INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Dense Estimation and Object-Oriented Segmentation of the Optical Flow with Robust Techniques

Etienne Mémin, Patrick Pérez, Denis Machecourt

N° 2836

Mars 1996

THÈME 3

ISSN 0249-6399

INRIA
RENNES

Dense Estimation and Object-Oriented Segmentation of the Optical Flow with Robust Techniques

Etienne Mémin, Patrick Pérez, Denis Machecourt

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Temis

Rapport de recherche n ° 2836 — Mars 1996 — 41 pages

Abstract: In this paper we address the intricate issue of recovering and segmenting the apparent velocity field between consecutive frames of an image sequence. As for motion estimation, we design a global cost functional to be minimized among the set of possible velocity fields. In the same spirit as the energy function proposed by Black *et al*, we consider two terms, both involving a *robust M-estimator*. The first one reinforces the fragile modeling of the *optical flow constraint equation*, while the second (*a priori*) term incorporates a *discontinuity preserving* smoothness constraint. A multiresolution formulation of this differential estimation method aims at accessing long range displacements in a coarse-to-fine incremental way. As for the minimization associated with robust estimators, we define a very efficient deterministic *multigrid relaxation* algorithm which converges fast toward estimates of good quality. Besides, the resulting iterative estimator is able to produce, at very low cost, "crude" estimates revealing the large discontinuity structures of the apparent motion field. This nice by-product of the robust multigrid estimation, encouraged us to incorporate a motion-based segmentation process. To do that, we propose an extension of the model by attaching to it an "object-oriented" segmentation device based on an interacting deformable closed curve. Different kinds of curve family equipped with different kinds of *a priori* energy function can be easily supported within this framework. We illustrate this by choosing two extreme cases in terms of curve's parametrization.

Experimental results on real world sequences are presented

Key-words: optical flow, multiresolution estimation, robust statistics, discontinuity preservation, global minimization, multigrid algorithms, iteratively reweighted least squares, object-oriented motion segmentation, closed segmenting curve

(Résumé : tsvp)

Unit´e de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)
T´el´ephone : (33) 99 84 71 00 – T´el´ecopie : (33) 99 84 71 71

Résumé : L'estimation du champ des vitesses apparentes dans une séquence d'images, et la segmentation basée mouvement des images de ladite séquence sont deux importants problèmes de vision dynamique qui se révèlent être étroitement imbriqués. C'est à ces problèmes que nous nous intéressons ici. Concernant l'estimation des vitesses apparentes, une fonction de coût à minimiser est définie. Inspirée des travaux de Black *et al.*, elle comporte deux termes basés chacun sur un M-estimateur robuste. Le premier terme permet de mettre en œuvre de façon robuste l'*équation de contrainte du mouvement apparent*. Le deuxième permet d'introduire un lissage régularisant qui sache *preserver les discontinuités* du flot optique. Le modèle différentiel ainsi obtenu est incorporé dans un cadre multirésolution pour permettre une estimation des grands déplacements. Ceci débouche sur un ensemble de problèmes de minimisation globale non-linéraire. Pour les traiter, une méthode *multigrille* très performante est mise au point. L'un de ses attraits, est qu'elle fournit à faible coût de calcul, des estimées "grossières" révélant les grandes structures de discontinuité du champs des vitesses apparentes. Ceci nous a encouragés à réaliser, *conjointement*, une segmentation au sens du mouvement qui tire partie de ces informations, tout en améliorant l'estimée finale du flot optique et en fournissant des renseignements compactes sur le contenu dynamique de la scène. Nous proposons ainsi une extension du précédent modèle en le dotant d'un outil de segmentation "basé-objet". Le principe est celui de la manipulation jointe de courbes déformables "segmentatrices", intéragissant avec l'estimation du flot optique, et destinées à délimiter le contour d'entités de mouvement. Différents types de courbes (en termes de paramétrisation de plus ou moins grande dimension) associés à différents types de connaissances *a priori* (sur la forme moyenne des objets recherchés par exemple) peuvent être aisément mis en œuvre dans le cadre de ce modèle couplé. Nous illustrons ceci en choisissant deux cas extrêmes de paramétrisations.

Des résultats expérimentaux obtenus sur séquences réelles sont présentés.

Mots-clé : flot optique, estimation multirésolution, statistique robuste, préservation des discontinuités, optimisation globale, algorithmes multigrilles, moindres carrés pondérés itérés, segmentation basée object au sens du mouvement, courbe fermée segmentatrice

Notations

$f(s,t):\ f(t)=\{f(s,t), s\in S\}$.	Luminance at pixel s at time t; image at time t.
$\boldsymbol{w}_s=[u_s,v_s]^T$	Apparent velocity at site s at time t.
$\boldsymbol{w}=\{\boldsymbol{w}_s, s\in S\}\in\Omega$	Apparent velocity field at time t.
$\boldsymbol{\nabla} f=[f_x,f_y]^T$	Spatial gradient of the luminance function.
f_t	Temporal partial derivative of luminance function.
$\nu;\ \nu(s)$	First-order neighborhood system; neighborhood of site s.
$\mathcal{C}$	Set of neighboring site pairs of S.
$k=0,\ldots,N$	Resolution level.
$l=0,\ldots,K$	Grid level associated with some resolution level.
$f^k(t)=\{f^k(s,t), s\in S^k\}$	Luminance function at resolution k.
$\boldsymbol{w}^k$	Compensating velocity field at resolution k.
$\mathsf{d}\boldsymbol{w}^k$	Increment velocity field to estimate at resolution k.
$f_t^k(s,t,\boldsymbol{w}_s^k)$	Displaced frame difference.
$H^k(\mathsf{d}\boldsymbol{w}^k; f^k,\boldsymbol{w}^k)$	Global energy function to be minimized w.r.t $\mathsf{d}\boldsymbol{w}^k$.
$\alpha>0$	Regularization parameter.
$\rho_1,\ \rho_2$	Robust M-estimator associated with the data and the prior.
σ_1,σ_2	Scale parameters associated with ρ_1 and ρ_2 respectively.
$\mathcal{H}^k(\mathsf{d}\boldsymbol{w}^k,\delta^k,\beta^k; f^k,\boldsymbol{w}^k)$	Dual global energy function with auxiliary variables.
$\delta^k=\{\delta_s^k, s\in S^k\}$	Data auxiliary variables.
$\beta^k=\{\beta_{sr}^k, <s,r>\in\mathcal{C}^k\}$	Discontinuity auxiliary variables.
$\phi_1(v)=\rho_1(\sqrt{v}),\ \phi_2(v)=\rho_2(\sqrt{v})$	Concave functions associated with ρ_1 and ρ_2.
M_1, M_2	Limits of $\frac{\rho_1'(u)}{2u}$ and $\frac{\rho_2'(u)}{2u}$, as u tends to 0.
$\mathsf{d}\boldsymbol{w}^{k,l}$	Reduced increment field lying on reduced grid $S^{k,l}$, and associated to piece-wise constant increment field $\Phi^{k,l}(\mathsf{d}\boldsymbol{w}^{k,l})\in\Omega^k$.
$\mathcal{H}^{k,l}(\mathsf{d}\boldsymbol{w}^{k,l},\delta^k,\beta^k; f^k,\boldsymbol{w}^k)$	Global energy function for grid level l.
$R\subset S$	Single object segment defined by a closed curve.
$\mathcal{C}_{\partial R}=\{<s,r>\in R\times\bar{R}\}$	Set of neighboring site pairs straddling the segment and its complement.
$\mathcal{C}_R\triangleq\{<s,r>\subset R\}$	Set of neighboring site pairs included in the segment.
$\mathbf{1}_R(s,r),\mathbf{1}_{\partial R}(s,r)$	Characteristic functions of $\mathcal{C}_R$ and $\mathcal{C}_{\partial R}$.
$\mathbb{H}(\mathsf{d}\boldsymbol{w},\delta,\beta,R; f,\boldsymbol{w})$	Global energy function of the augmented estimation/segmentation model
$E_{prior}(R)$	Energetic prior about the admissible segments.
$\Delta\boldsymbol{w}_{sr}=(\boldsymbol{w}_s+\mathsf{d}\boldsymbol{w}_s)-(\boldsymbol{w}_r+\mathsf{d}\boldsymbol{w}_r)$	Pair-wise velocity change.
$\hat{\beta}_{sr}^0\triangleq\frac{1}{M_2}\phi_2'[\|\Delta\boldsymbol{w}_{sr}\|^2]$	Segmentation-free optimal discontinuity weight .

1 Introduction

Many tasks in computer vision and image analysis can be expressed as global optimization problems [12, 27, 31]. The general issue is to find the global minimum of an objective (also called *energy*) function which describes the interaction between the different variables modeling the image features in a given problem.

Two kinds of variables are generally considered: *observation variables* which correspond to the representation of the observed data and *hidden variables* which are the information to be extracted from the original images [1]. Energy functions generally involve two components. A first component expresses the interaction between the unknown variables and the observations, while the other one captures some kind of prior knowledge about the set of unknown variables. This later ingredient is often a mere *regularization* term which only encodes a dim prior, but whose essential role is to cure the ill-posed nature of the problem at hand (to some extent, it guarantees the existence and the uniqueness of a consistent solution which continuously depends on the data). To keep the energy-based model tractable, the cost function is often decomposed as a sum of local interaction functions associated with a neighborhood system. Standard regularization approaches [39] as well as Markov Random Field (MRF)-based image analysis [19] or Minimum Description Length principle [31] lead to the minimization of such global energy functions. With Markov Random Field-based models for instance, a probability distribution is defined. Its use within a Bayesian estimation framework naturally leads to the minimization of an energy function [19].

Within this energy minimization framework, we address here the particular problem of *optical flow estimation*, and its possible association with some *optical flow-based segmentation*. The motion estimator that we present here belongs to the class of differential estimation methods. These methods rely on the well known *optical flow constraint* (ofce) which links within a single equation the spatio-temporal derivatives of the luminance function and the motion field to estimate. Unfortunately, this constraint is very sensitive to noise and behaves badly when spatial or temporal variations of the luminance function are too large. An other source of concern arises from the smoothness regularizing prior which is usually associated to these models. The frontiers which demarcate the different apparent motions coexisting within the same scene are ignored by a blind smoothing. This deficiency results in a bad estimation nearby these border lines. A great deal of studies has been dedicated to this specific problem of discontinuity preserving regularization in computing optical flow (and in computer vision in general). See for instance [10, 13, 15, 24, 37, 40, 41].

Adopting a more global viewpoint, M. J. Black points out in [5] that all the different problems we have just evoked can be seen as "deviations" from the data model and from the prior model respectively. Though different in nature, they can hopefully be located and treated

[1] Any occurrence of this set of hidden variables will be often called "configuration" throughout the paper.

within a unified framework: the one offered by the *robust statistics* whose original aim is the estimation of models in presence of many deviating occurrences among the data [28].

Nevertheless, the introduction of robust estimators in energy-based image applications leads most of the time to a global non-linear minimization in presence of numerous local minima. Even within a *multiresolution* formulation of the problem (which is almost inescapable in case of long range motions to be estimated), one has to deal with a sequence of global optimization problems which remain tricky.

To avoid the use of greedy stochastic algorithms [19], some authors [5, 31] have proposed to take benefits from the "scale" parameter involved in standard robust estimators. A proper and progressive tuning of this parameter allows to define a minimization strategy similar in spirit to continuation methods such as the "Graduated Non Convexity" algorithm proposed by Blake and Zisserman [9].

For the same purpose (i.e., definition of an efficient deterministic algorithm to deal with the global optimization problem at hand), we propose here to extend the *multigrid* relaxation method presented in [25]. This method consists in the minimization of the energy function through a hierarchy of nested subspaces of the whole configuration space. These subspaces correspond to configurations constrained to be piecewise constant over smaller and smaller pixel subsets. This is equivalent to handling velocity fields which lie on reduced grids, whereas observations are still viewed at the current resolution.

It is well known that the estimation of the optical flow in the one hand, and the segmentation of the images with respect to the apparent motion in the other hand, are two important issues of motion analysis which should help each other to provide better results, as well as richer information. Following this philosophy, we propose a connection of the optical flow estimator under concern with an object-oriented motion segmentation. This is addressed within an *augmented objective function* which involves one or several segmenting closed-curves.

From the point of view of the optical flow estimation this curve is used to hopefully drive and structure the apparition of spatial discontinuities, thus improving the quality of the estimation. Conversely, the segmenting curve, which provides interesting information of its own, is driven by the spatial discontinuities of the velocity field under estimation. Without much pain, the efficient multigrid structure can be adapted to this joint model.

The paper is organized as follows. In Section 2, we present the very classical gradient-based optical flow estimation, relying on an infinitesimal expansion of the brightness constancy assumption. The notational basis of the paper is set up at that time. From this standard framework, an energy-based multiresolution model for a robust and discontinuity preserving estimation of the apparent motion can be developed, as explained in Section 3. The resulting hierarchical robust model is rewritten by using a dual quadratic formulation involving various auxiliary variables. The global optimization issue is then addressed in Section 4. A deterministic multigrid algorithm involving iteratively reweighted least squares estimation is proposed. In Section 5, we explore a way of coupling the motion estimation process with a simultaneous

segmentation of the current optical flow. We especially show how a segmenting closed curve can interact with the discontinuity auxiliary variables to extract moving entities while enhancing the optical flow estimation nearby. In Section 6, experimental results of optical flow estimation and object-oriented motion segmentation are displayed on real-world sequences.

2 Optical flow estimation

Let $\boldsymbol{w} = \{\boldsymbol{w}_s, s \in S\} \in \Omega$ and $f(t) = \{f(s,t), s \in S\}$ be respectively the unknown velocity field at time t and the luminance function at time t, both defined on the rectangular pixel lattice S. The apparent velocity at location s of the image plane is $\boldsymbol{w}_s = [u_s, v_s]^T$, and the configuration space to explore, Ω, is some bounded subset of $(\mathbb{R} \times \mathbb{R})^S$.

Assuming as in most approaches that the luminance of a given "physical point" does not change much between times t and $t + dt$, one gets:

$$f(s + dt\boldsymbol{w}_s, t + dt) - f(s,t) \approx 0. \tag{1}$$

This equation is non-linear with respect to the unknown velocity vector. In order to linearize it, a first order expansion is usually computed around (s,t). It leads to:

$$\varepsilon_s \triangleq \boldsymbol{\nabla} f(s,t).\boldsymbol{w}_s + f_t(s,t) \approx 0, \tag{2}$$

where " . " denotes the standard Euclidean inner product, f_t stands for the temporal partial derivative of luminance f and $\boldsymbol{\nabla} f = [f_x, f_y]^T$ is its spatial gradient [2]. This equation is known as the *optical flow constraint equation* (OFCE) and relates spatio-temporal luminance changes to the unknown velocity field $\boldsymbol{w}$. This single equation does not allow to recover the two motion components at each pixel location, and it is very sensitive to noise. For these reasons, a contextual modeling such as *smoothness* properties in regularization theory [39] or Gibbs *prior* distribution on the set of velocity fields, must be added.

Within Bayesian framework, for example, the $\varepsilon_s, s \in S$, are usually assumed to be independent realizations of some Gaussian white noise, while $\boldsymbol{w}$ is *a priori* a "smooth" Gauss-MRF with respect to some neighborhood system on S. Hereafter, we will only consider the 4-neighborhood system (denoted ν) for which any site apart from those on the image "border" has a set $\nu(s)$ of 4 (nearest) neighbors. In association with this neighborhood system, the considered smoothness is naturally of the first order [3]. Estimating the best field according to the *Maximum A Posteriori* (MAP) Bayesian criterion leads to the global optimization problem:

$$\widehat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w} \in \Omega} H(\boldsymbol{w}; f), \tag{3}$$

[2] As a convenience, the brightness function f is assumed to be differentiable with respect to space and time (as it was implicitly assumed continuous in space in (1)). As for practical issues, the spatial gradient of f is computed with the optimal derivative filters described in [45], while the temporal derivative is classically approximated by smoothing the frame-to-frame difference.

[3] Note that smoothness of higher orders may be considered, in association with neighborhood systems of higher orders, as described in [18].

with

$$H(\boldsymbol{w}; f) \triangleq \sum_{s \in S} [\boldsymbol{\nabla} f(s,t).\boldsymbol{w}_s + f_t(s,t)]^2 + \alpha \sum_{<s,r> \in \mathcal{C}} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 , \quad (4)$$

where $\mathcal{C}$ is the set of neighboring site pairs (with respect to the neighborhood system ν): $<s,r> \in \mathcal{C} \Leftrightarrow r \in \nu(s) \Leftrightarrow s \in \nu(r)$. The design of such a cost functional to be minimized can also be done outside any probabilistic framework, as in [27]. Anyway, the parameter $\alpha > 0$ turns out to control the balance between the smoothness constraint and the global adequacy to the OFCE. This energy function is quadratic with respect to the unknown vector field. Its minimization leads to a set of linear equations which can be processed through the following Gauss-Seidel iterative scheme:

$$\boldsymbol{w}_s^{(n+1)} = \overline{\boldsymbol{w}}_s^{(n)} - \frac{\left[\boldsymbol{\nabla} f(s,t).\boldsymbol{w}_s^{(n)} + f_t(s,t)\right]}{|\nu(s)|\alpha + \|\boldsymbol{\nabla} f(s,t)\|^2} \boldsymbol{\nabla} f(s,t), \quad \text{with } \overline{\boldsymbol{w}}_s^{(n)} \triangleq \frac{1}{|\nu(s)|} \sum_{r \in \nu(s)} \boldsymbol{w}_r^{(n)}, \quad (5)$$

where sites of S are visited iteratively, and s stands above for the site visited at step $n+1$. The new velocity vector at site s is then obtained by averaging those of the neighbors (smoothing propagation) and subtracting a data-driven ratio related to the former velocity.

This iterative scheme is similar to Horn and Shunck's one [27]. Note that the MRF formulation leads naturally to a discrete modeling of the problem whereas it is continuous in [27] (solutions being discretized afterwards).

The drawbacks of this quadratic gradient-based model are well known:

1. Due to its differential nature, the OFCE is not valid in case of large displacements;
2. For the same kind of reasons, the OFCE is not valid in areas exhibiting very large (spatial or temporal) luminance variations: occlusions involving appearance or disappearance of scene's portions, sharp edges, specular reflection patches...
3. The quadratic regularization results in over-smoothed estimates which do not respect the natural discontinuities usually present in apparent motion fields.

We see in the coming section how these drawbacks can be partially overcome.

3 Robust multiresolution optical flow estimation

3.1 Multiresolution setup

To address the first problem listed above (namely, the recovery of large displacements), a standard technique consists in using a multiresolution model along with a *coarse-to-fine incremental* estimation [5, 16, 24]. To this end, for each instant t of the sequence, a pyramid of images $f^k(t) = \{f^k(s,t), s \in S^k\}$, $k = 0 \ldots N$, is derived from the original frame $f(t)$ by

successive Gaussian smoothings and regular subsamplings by a factor 2 in each direction [4]. The resolution index k, which will be used as a superscript throughout the paper, spans from 0 for the finest resolution, to N for the coarsest one. The created pyramidal structure is then used to estimate *incrementally* the velocity field: long range components are estimated at coarse scales and then progressively refined through finer resolutions. More precisely, a first estimate is searched at the coarsest resolution where the domain of validity of the OFCE is hopefully larger (due to the reduction of noise by smoothing, and the reduction of motion magnitude by subsampling). This crude estimate is then refined step by step: at resolution k, an *increment velocity field* $\mathsf{d}\boldsymbol{w}^k$ in $\Omega^k \subset (\mathbb{R}\times\mathbb{R})^{S_k}$ is estimated around the last estimate $\boldsymbol{w}^k$ provided by the previous resolution level $(k+1)$. At that stage, the brightness constancy assumption becomes:

$$f^k(s+dt[\boldsymbol{w}_s^k+\mathsf{d}\boldsymbol{w}_s^k],t+dt)-f^k(s,t)\approx 0. \tag{6}$$

A first order expansion around $(s+dt\boldsymbol{w}_s^k,t+dt)$ yields an OFCE *displaced according to the last field estimate* $\boldsymbol{w}^k$:

$$\boldsymbol{\nabla} f^k(s+dt\boldsymbol{w}_s^k,t+dt).\mathsf{d}\boldsymbol{w}_s^k+\underbrace{\left[f^k(s+dt\boldsymbol{w}_s^k,t+dt)-f^k(s,t)\right]/dt}_{\triangleq f_t^k(s,t,w_s^k)}\approx 0. \tag{7}$$

This equation is formally similar to the regular OFCE. Especially, it is also linear with respect to the unknown variables. As for the differences, the spatial gradient is in this case computed in the second frame (i.e, a time $t+dt$), at the displaced location $s+dt\boldsymbol{w}_s^k$, and the temporal derivative is now replaced by a *displaced frame difference*, here denoted $f_t^k(s,t,\boldsymbol{w}_s^k)$.

3.2 Robust formulation

The two others sources of concern listed at the end of §2, can be addressed by introducing *robust estimators* in the energy formulation at each resolution. By doing so, one wants to cope with large *deviations* from the data model (here, the displaced OFCEs (7)) and from the prior model (here, the first order smoothing prior), which are very likely to occur. In robust statistics vocabulary these deviations are called *outliers*. In our case, they are therefore of two types: outliers attached to the presence of spatial discontinuities in the velocity field to estimate; and outliers associated with areas where the OFCE is not valid at all. Considering energy function (4), which consists of a data model part (sum of site-wise data model deviation) and a smoothing part (sum of local smoothness measures), it seems natural to embed robust M-estimators within both terms, as proposed by Black *et al.* [5, 7]. Namely, at each resolution level k, we consider the following energy function of $\mathsf{d}\boldsymbol{w}^k$ given $\boldsymbol{w}^k$ and f^k:

$$H^k(\mathsf{d}\boldsymbol{w}^k;f^k,\boldsymbol{w}^k)\triangleq H_1^k(\mathsf{d}\boldsymbol{w}^k;f^k,\boldsymbol{w}^k)+\alpha H_2^k(\mathsf{d}\boldsymbol{w}^k;\boldsymbol{w}^k), \tag{8}$$

[4] The grid S^k is 4^k times smaller than S.

with:

$$H_1^k(\mathsf{d}\boldsymbol{w}^k; f^k, \boldsymbol{w}^k) \triangleq \sum_{s \in S^k} \rho_1 \left[\boldsymbol{\nabla} f^k(s + dt\boldsymbol{w}_s^k, t + dt).\mathsf{d}\boldsymbol{w}_s^k + f_t^k(s, t, \boldsymbol{w}_s^k) \right], \quad (9)$$

$$H_2^k(\mathsf{d}\boldsymbol{w}^k; \boldsymbol{w}^k) \triangleq \sum_{<s,r> \in \mathcal{C}^k} \rho_2 \left[\|(\boldsymbol{w}_s^k + \mathsf{d}\boldsymbol{w}_s^k) - (\boldsymbol{w}_r^k + \mathsf{d}\boldsymbol{w}_r^k)\| \right], \quad (10)$$

where $\mathcal{C}^k$ is the set of neighboring site pairs lying on grid S^k equipped with the 4-neighborhood system and functions ρ_1 and ρ_2 are standard robust M-estimators, involving some scale parameters, σ_1 and σ_2 respectively. Figure 1 pictures this multiresolution energy setup.

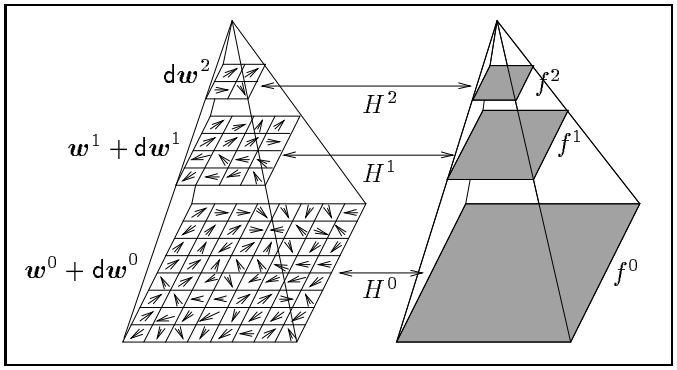

Figure 1: Structure of the multiresolution model ($N = 2$).

As for the coarse-to-fine incremental aspect, the energy functions have to be minimized with respect to $\mathsf{d}\boldsymbol{w}^k$, from $k = N$ (with $\boldsymbol{w}^N \equiv 0$), to $k = 0$, the current field $\boldsymbol{w}^k$ to refine being a "projection" (e.g., obtained by duplication, or bilinear interpolation) of the estimate $\widehat{\boldsymbol{w}}^{k+1}$ obtained at the previous level. The minimization of $H^k(\mathsf{d}\boldsymbol{w}^k; f, \boldsymbol{w}^k)$ yields $\widehat{\mathsf{d}\boldsymbol{w}}^k$. The sum estimate $\widehat{\boldsymbol{w}}^k = \boldsymbol{w}^k + \widehat{\mathsf{d}\boldsymbol{w}}^k$ is then refined at resolution $k-1$, and so forth. At the finest resolution $k = 0$, the final estimate $\boldsymbol{w}^0 + \widehat{\mathsf{d}\boldsymbol{w}}^0$ is produced. Figure 2 depicts the whole descending process.

Let's now turn our attention to the robust M-estimators. Roughly speaking, they are continuous even "cost" functions, increasing on $\mathbb{R}^+$, and which penalize large "residual" values less drastically than quadratic functions do. More precisely, their derivatives, called "influence functions" in robust statistics [28], have finite limits at infinity (usually zero). In the model currently discussed, this characteristic makes them robust to data model outliers (for ρ_1) or to spatial discontinuities of the apparent motion field (for ρ_2).

To give an insight into robust estimation as well as a practical way of handling it, it is fruitful transforming its use in terms of dual optimization problem involving auxiliary variables [1, 5, 18]:

PROPOSITION. *Let ρ be a real-valued continuously differentiable even function such that:*

- *ρ is strictly increasing on $\mathbb{R}^+$;*
- *$\phi(v) \triangleq \rho(\sqrt{v})$ is strictly concave on $\mathbb{R}^+$;*
- *$\lim_{u \to +\infty} \rho'(u) < +\infty$.*

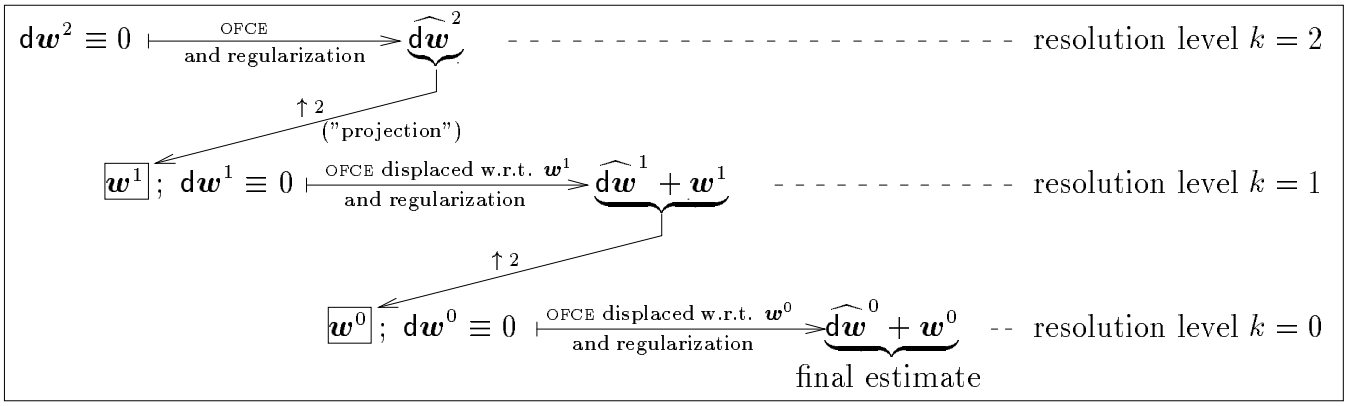

Figure 2: Synopsis of the incremental coarse-to-fine optical flow estimation on three levels ($N = 2$).

There exits a function ψ, continuously differentiable on $(0, M]$ (with $M \triangleq \lim_{v \to 0^+} \phi'(v) \in (0, +\infty]$), such that:

$$\forall u \in \mathbb{R}^+,\ \rho(u) = \min_{z \in (0,M]} zu^2 + \psi(z). \tag{11}$$

This means that the graph of ρ is the inferior envelope of a family of parabolas continuously indexed by $z \in (0, M]$. Equivalently, the graph of ϕ is the inferior envelope of a family of lines indexed in the same way. Figure 3 illustrates these properties in case of Leclerc's estimator [31].

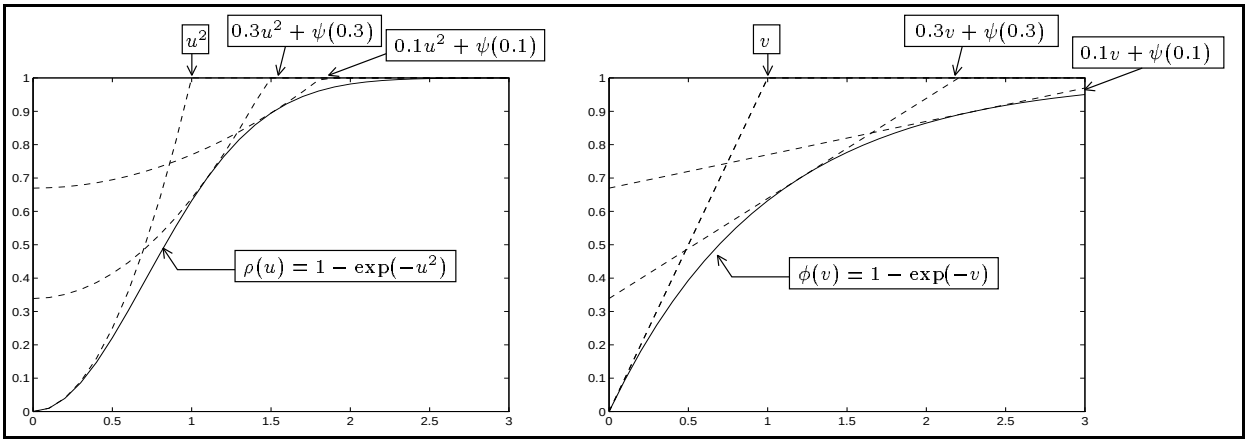

Figure 3: Graph of ρ (resp. ϕ) as the inferior envelope of a family of parabolas (resp. lines). Here $\rho(u) = 1 - \exp(-u^2)$, and therefore, $\phi(v) = 1 - \exp(-v)$, $\psi(z) = 1 + z(\ln z - 1)$, $M = 1$.

The minimum in (11) is given by:

$$\operatorname*{argmin}_{z \in (0,M]} zu^2 + \psi(z) = \frac{\rho'(u)}{2u} = \phi'(u^2). \tag{12}$$

In Appendix A, we give a sketch of the proof inspired from [18] (other formulations are given in [1, 5]). This proposition allows to turn any multidimensional minimization problem of the form

$$\text{find } \arg\min_x \sum_i \rho[g_i(x)], \tag{13}$$

into a dual minimization problem:

$$\text{find } \arg\min_{x,z} \sum_i [z_i g_i(x)^2 + \psi(z_i)], \tag{14}$$

involving *auxiliary variables* z_is continuously lying in $(0, M]$. As a matter of fact, sums in (13) and (14) have the same global minimum in x. Note that in practice, each function g_i only depends on a few components of x (one or two in our problem, as we shall see).

The minimization in the dual problem is usually lead alternatively with respect to x in the one hand, and with respect to the z_is in the other hand:

- the z_is being "frozen", the minimization in x becomes:

$$\arg\min_x \sum_i [z_i g_i(x)^2 + \psi(z_i)] = \arg\min_x \sum_i z_i g_i(x)^2 = \hat{x}(z).$$

  If g_is are affine forms, one has to face a standard *weighted least squares* problem equivalent to the resolution of some linear system.

- x being frozen, the minimization with respect to auxiliary variables yields

$$\hat{z}_i(x) = \frac{\rho'[g_i(x)]}{2g_i(x)} = \phi'[g_i(x)^2], \tag{15}$$

  according to (12).

In case g_is are affine, the auxiliary variables can be interpreted as *weights* which are adaptatively updated from time to time according to the evolution of the main estimation process (i.e., w.r.t. x). The whole alternate procedure constitutes an *iteratively reweighted least squares* estimation [26].

The assumptions about ρ mean that ϕ' strictly decreases from $\mathbb{R}^+$ into $(0, M]$: when the ith "residual" $g_i(x)^2$ gets larger, the corresponding optimal weight $\hat{z}_i$ gets smaller and smaller, providing the robustness of the estimator.

Figure 4 shows three standard robust estimators and their associated optimal weight functions (called "interactions" in [33]). For practical reasons, it might be advantageous to have weights which belong to $[0, 1]$. If $M < \infty$, which is generally the case (though not always; see [18] for instance), one can consider a rescaled weight $z^* \triangleq \frac{z}{M}$. The minimization formulation of ρ becomes:

$$\rho(u) = \min_{z^* \in (0,1]} M z^* u^2 + \psi(M z^*). \tag{16}$$

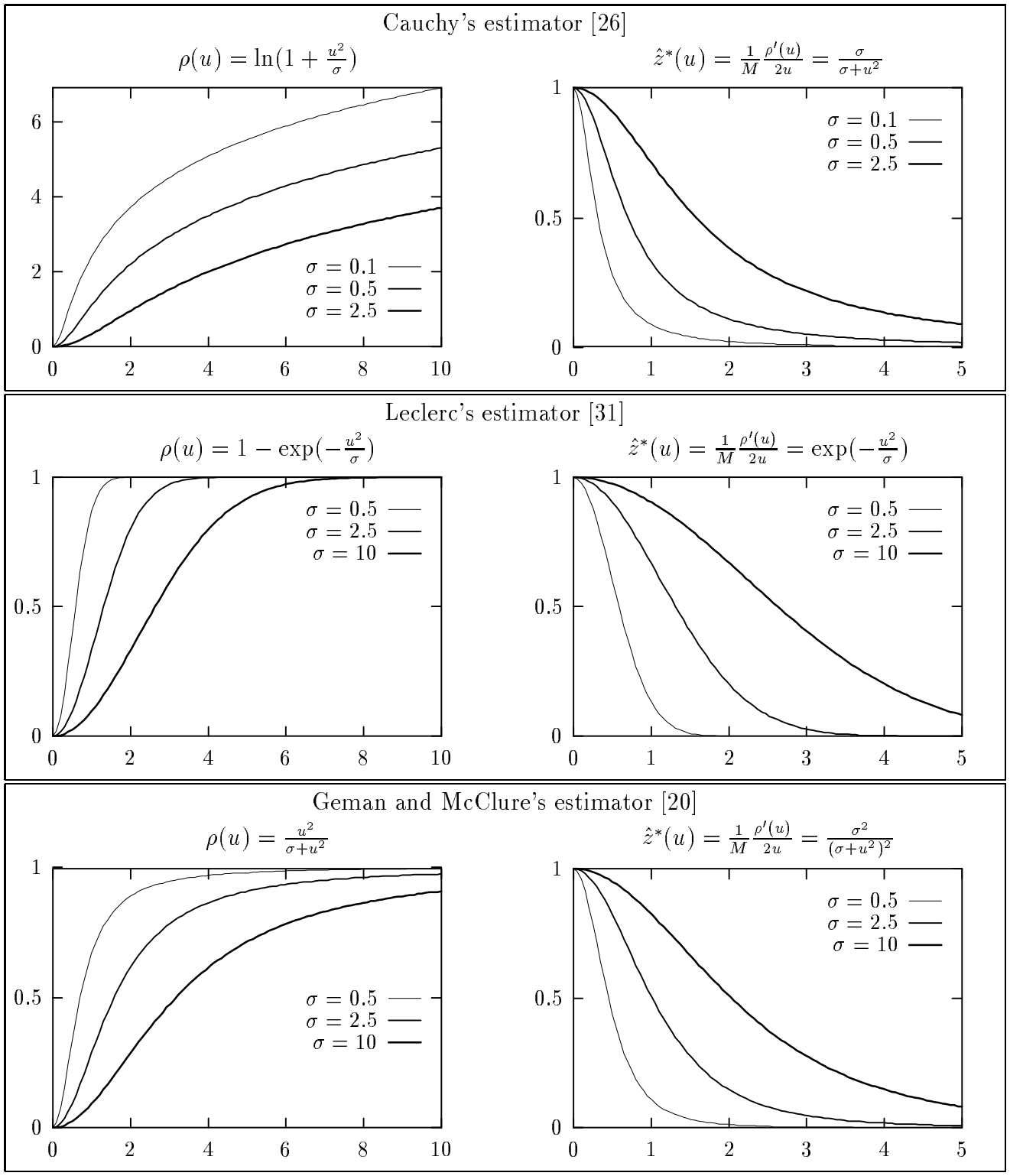

Figure 4: Three robust estimators and their associated optimal weight functions. For each of them, $M = \lim_{u\to 0^+} \frac{\rho'(u)}{2u} = \frac{1}{\sigma}$.

The estimators we actually chose (see §5) in our experiments allow this rescaling. Consequently, all auxiliary variables considered hereafter are meant to be the rescaled ones.

The advantages of transforming a robust estimator-based energy in its auxiliary variable formulation are twofold:

(a) It allows the model to be extended by adding interactions among auxiliary variables (e.g., to capture some "geometric" *a priori* knowledge, either local or more global, on the discontinuity configurations). In this paper, we shall present an extension of the model in this direction which aims at coupling the robust discontinuity preserving motion estimation with an object-oriented motion segmentation.

(b) As we previously mentioned, the minimization of the resulting energy function when done in an alternate way, leads to an *iteratively reweighted least squares* minimization [26] which can be performed with some standard quadratic minimization tools.

In our case the weights are of two natures: (a) *data outliers auxiliary variables* (related to the dual formulation of H_1^k), and (b) *discontinuity auxiliary variables* lying on the dual grid of S^k (provided by the dual formulation of H_2^k). The first set of weights, denoted by $\delta^k = \{\delta_s^k, s \in S^k\}$, allows to attenuate the effect of data for which the OFCE is violated. The second one, denoted by $\beta^k = \{\beta_{sr}^k, <s,r> \in \mathcal{C}^k\}$, prevents from over-smoothing in locations obviously exhibiting significant velocity discontinuities. The estimation at resolution level k is now expressed as the global minimization of:

$$\mathcal{H}^k(\mathsf{d}\boldsymbol{w}^k, \delta^k, \beta^k; f^k, \boldsymbol{w}^k) \triangleq \mathcal{H}_1^k(\mathsf{d}\boldsymbol{w}^k, \delta^k; f^k, \boldsymbol{w}^k) + \alpha \mathcal{H}_2^k(\mathsf{d}\boldsymbol{w}^k, \beta^k; \boldsymbol{w}^k), \tag{17}$$

with:

$$\mathcal{H}_1^k(\mathsf{d}\boldsymbol{w}^k, \delta^k; f^k, \boldsymbol{w}^k) \triangleq \sum_{s \in S^k} \left[M_1 \delta_s^k \left[\boldsymbol{\nabla} f^k(s + dt\boldsymbol{w}_s^k, t + dt).\mathsf{d}\boldsymbol{w}_s^k + f_t^k(s, t, \boldsymbol{w}_s^k) \right]^2 + \psi_1(M_1 \delta_s^k) \right] \tag{18}$$

$$\mathcal{H}_2^k(\mathsf{d}\boldsymbol{w}^k, \beta^k; \boldsymbol{w}^k) \triangleq \sum_{<s,r> \in \mathcal{C}^k} \left[M_2 \beta_{sr}^k \| (\boldsymbol{w}_s^k + \mathsf{d}\boldsymbol{w}_s^k) - (\boldsymbol{w}_r^k + \mathsf{d}\boldsymbol{w}_r^k) \|^2 + \psi_2(M_2 \beta_{sr}^k) \right]. \tag{19}$$

However, the underlying energy function H^k being non-convex with respect to the unknown variables of interest, we still have to deal with a tough optimization problem, in spite of the reformulation. In particular, the alternate minimization procedure is not guaranteed to reach the global minimum, even though each step is constituted of an exact minimization (but with respect to only a subset of variables). Actually, only a "local" minimum depending on the initialization is reached. For this reason, we designed an extension of the *multigrid method* proposed in [25] which hopefully converges fast without getting stuck in high local minima.

4 Multigrid deterministic optimization

To efficiently cope with the global optimization problem at resolution k, we design a hierarchical "constrained" exploration of the configuration space Ω^k: the optimization is lead through a sequence of nested configuration subspaces:

$$\Omega^{k,L} \subset \Omega^{k,L-1} \subset \cdots \subset \Omega^{k,1} \subset \Omega^{k,0} = \Omega^k, \tag{20}$$

where $\Omega^{k,l}$ is the set of increment fields which are piecewise constant according to a $2^l \times 2^l$-block partition of grid S^k. Denote $\mathcal{B}^{k,l} \triangleq \{\mathcal{B}_n^{k,l},\ n = 1 \dots N_{k,l}\}$ this partition, the number of blocks being $N_{k,l} = |S^k|/4^l = |S|/4^{l+k}$. Each constrained configuration of $\Omega^{k,l}$ is "equivalent" to a *reduced increment field* $\mathsf{d}\boldsymbol{w}^{k,l}$ lying on the grid $S^{k,l} \triangleq \{1, \dots, N_{k,l}\}$ associated with $\mathcal{B}^{k,l}$. Let $\Gamma^{k,l}$ be the set of such reduced configurations and let $\Phi^{k,l}$ be the point-to-point mapping from $\Gamma^{k,l}$ into $\Omega^{k,l}$ (Fig. 5).

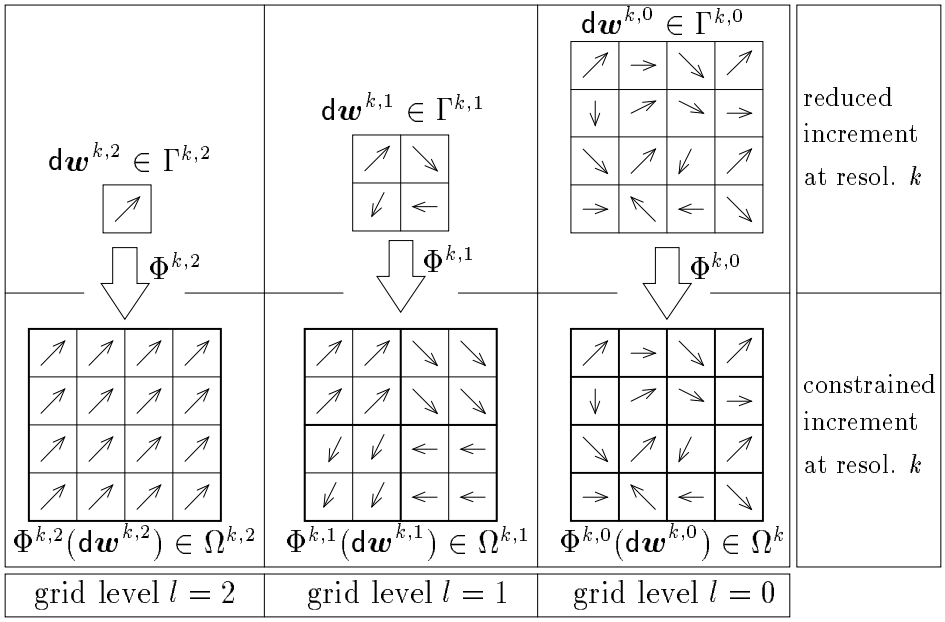

Figure 5: Three levels of constrained increment fields at resolution level k, and reduced increment fields associated to them ($L = 2$).

Constrained optimization in $\Omega^{k,l}$ is then equivalent to the minimization of the new energy function:

$$\mathcal{H}^{k,l}(\mathsf{d}\boldsymbol{w}^{k,l}, \delta^k, \beta^k; f^k, \boldsymbol{w}^k) \triangleq \mathcal{H}^k(\Phi^{k,l}(\mathsf{d}\boldsymbol{w}^{k,l}), \delta^k, \beta^k; f^k, \boldsymbol{w}^k) \tag{21}$$

defined over $\Gamma^{k,l}$, whereas the auxiliary variables, the data, and the field to be refined remain the same (i.e., defined on grid S^k) (Fig. 6).

At each resolution, we now have a cascade of optimization problems of reduced complexity:

$$\arg\min_{\mathsf{d}\boldsymbol{w}^{k,l}, \delta^k, \beta^k} \mathcal{H}^{k,l}(\mathsf{d}\boldsymbol{w}^{k,l}, \delta^k, \beta^k; f^k, \boldsymbol{w}^k),\ l = L \dots 0, \tag{22}$$

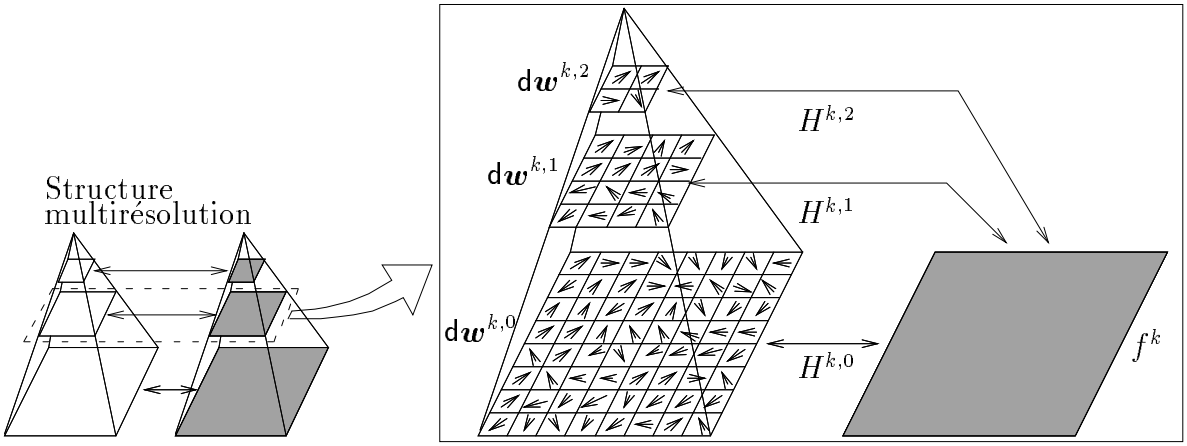

Figure 6: Multigrid structure for relaxation at level k ($L = 2$)

where $\mathsf{d}\boldsymbol{w}^{k,l} \in \Gamma^{k,l}$ lies on the reduced grid $S^{k,l}$, and auxiliary variables are attached to S^k, whatever the grid level l.

Each of these problems is processed in terms of iteratively reweighted least squares within a multigrid coarse-to-fine strategy: the final estimate at level $l+1$ has a natural image at level l (through $[\Phi^{k,l}]^{-1} \circ \Phi^{k,l+1}$) which is used as an initial configuration for the deterministic relaxation algorithm at that level. This procedure is repeated until the finest level $l = 0$ is reached (see [25] for further details).

4.1 Multigrid energy derivation

We now have to go into deeper details about the new multigrid functions $\mathcal{H}^{k,l}$. One can show that the energy at grid level l is composed of two terms similar to those of $\mathcal{H}^k$:

$$\mathcal{H}^{k,l}(\mathsf{d}\boldsymbol{w}^{k,l}, \delta^k, \beta^k; f^k, \boldsymbol{w}^k) = \mathcal{H}_1^{k,l}(\mathsf{d}\boldsymbol{w}^{k,l}, \delta^k; f^k, \boldsymbol{w}^k) + \alpha \mathcal{H}_2^{k,l}(\mathsf{d}\boldsymbol{w}^{k,l}, \beta^k; \boldsymbol{w}^k). \tag{23}$$

Let us develop these two energy terms. For sake of readability, we will omit the resolution superscript k in all expressions throughout the remainder of this section. *All computations will be meant to concern resolution level k.*

Data model adequation term: due to the pointwise interaction structure of $\mathcal{H}_1$, one can derive:

$$\begin{aligned}\mathcal{H}_1^l(\mathsf{d}\boldsymbol{w}^l, \delta; f, \boldsymbol{w}) &= \sum_{n \in S^l} \sum_{s \in \mathcal{B}_n^l} \left[M_1 \delta_s \left[\boldsymbol{\nabla} f(s + dt\boldsymbol{w}_s, t + dt).\mathsf{d}\boldsymbol{w}_n^l + f_t(s, t, \boldsymbol{w}_s) \right]^2 + \psi_1(M_1 \delta_s) \right] \\ &= \sum_{n \in S^l} \left[M_1 \sum_{s \in \mathcal{B}_n^l} \delta_s \left[\boldsymbol{\nabla} f(s + dt\boldsymbol{w}_s, t + dt).\mathsf{d}\boldsymbol{w}_n^l + f_t(s, t, \boldsymbol{w}_s) \right]^2 + \Psi_1^l(\delta_{\mathcal{B}_n^l}) \right] \end{aligned} \tag{24}$$

with:

$$\Psi_1^l(\delta_{\mathcal{B}_n^l}) \triangleq \sum_{s\in\mathcal{B}_n^l} \psi_1(M_1\delta_s). \tag{25}$$

Sites of block $\mathcal{B}_n^l$ being denoted by $s_1,\ldots,s_{4^l}$, define the following blockwise vectors and matrices:

$$\begin{aligned} F_{xy}^l(n,\boldsymbol{w})^T &\triangleq \left[\boldsymbol{\nabla} f(s_1+dt\boldsymbol{w}_{s_1},t+dt)\cdots\boldsymbol{\nabla} f(s_{4^l}+dt\boldsymbol{w}_{s_{4^l}},t+dt)\right] \\ &= \left[\begin{array}{c} \boldsymbol{F}_x^l(n,\boldsymbol{w})^T \\ \boldsymbol{F}_y^l(n,\boldsymbol{w})^T \end{array}\right], \end{aligned} \tag{26}$$

$$\boldsymbol{F}_t^l(n,\boldsymbol{w}) \triangleq [f_t(s_1,t,\boldsymbol{w}_{s_1})\cdots f_t(s_{4^l},t,\boldsymbol{w}_{s_{4^l}})]^T,$$

$$\Delta_n^l \triangleq \mathrm{diag}(\delta_{s_1},\ldots,\delta_{s_{4^l}}).$$

Since all δ_ss are positive, diagonal matrix Δ_n^l defines an inner product and a norm on $\mathbb{R}^{4^l}$. They will be denoted by $\langle\ |\ \rangle_n$ and $\|\ \|_n$ respectively:

$$\langle \boldsymbol{X}|\boldsymbol{Y}\rangle_n \triangleq \boldsymbol{X}^T\Delta_n^l\boldsymbol{Y}, \quad \|\boldsymbol{X}\|_n^2 \triangleq \langle \boldsymbol{X}|\boldsymbol{X}\rangle_n.$$

One finally gets the following compact expression:

$$\boxed{\mathcal{H}_1^l(\mathsf{d}\boldsymbol{w}^l,\delta;f,\boldsymbol{w}) = \sum_{n\in S^l}\left[M_1\|F_{xy}^l(n,\boldsymbol{w})\mathsf{d}\boldsymbol{w}_n^l + \boldsymbol{F}_t^l(n,\boldsymbol{w})\|_n^2 + \Psi_1^l(\delta_{\mathcal{B}_n^l})\right],} \tag{27}$$

which is *very similar* to the one of the "parent" energy $\mathcal{H}_1$ (18). For each block, one get a kind of blockwise optical flow constraint expression involving aggregated observations.

Smoothing term: let $\mathcal{C}_n^l \triangleq \{<s,r>\in\mathcal{C} :<s,r>\subset\mathcal{B}_n^l\}$ be the set of neighboring site pairs included in block $\mathcal{B}_n^l$ and $\mathcal{C}_{nm}^l \triangleq \{<s,r>\in\mathcal{C} : s\in\mathcal{B}_n^l, r\in\mathcal{B}_m^l\}$ the set of neighboring site pairs straddling blocks $\mathcal{B}_n^l$ and $\mathcal{B}_m^l$. Notice that in the 4-neighborhood case, $\mathcal{C}_n^l$ and $\mathcal{C}_{nm}^l$ site sets form a partition of $\mathcal{C}$ and reduced grid S^l turns out to be equipped with a 4-neighborhood system whose neighboring pair set will be denoted by $\mathcal{C}^l$. The smoothing term of $\mathcal{H}^l$ is:

$$\begin{aligned} \mathcal{H}_2^l(\mathsf{d}\boldsymbol{w}^l,\beta;\boldsymbol{w}) &= M_2\left[\sum_{n\in S^l}\sum_{<s,r>\in\mathcal{C}_n^l}\beta_{sr}\|\boldsymbol{w}_s-\boldsymbol{w}_r\|^2\right. \\ &\qquad \left. + \sum_{<n,m>\in\mathcal{C}^l}\sum_{<s,r>\in\mathcal{C}_{nm}^l}\beta_{sr}\|(\boldsymbol{w}_s+\mathsf{d}\boldsymbol{w}_n^l)-(\boldsymbol{w}_r+\mathsf{d}\boldsymbol{w}_m^l)\|^2\right] \\ &\quad + \sum_{<s,r>\in\mathcal{C}}\psi_2(M_2\beta_{sr}). \end{aligned} \tag{28}$$

Noting that

$$
\begin{aligned}
& M_2 \left[\sum_{n \in S^l} \sum_{<s,r> \in \mathcal{C}_n^l} \beta_{sr} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 + \sum_{<n,m> \in \mathcal{C}^l} \sum_{<s,r> \in \mathcal{C}_{nm}^l} \beta_{sr} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 \right] + \sum_{<s,r> \in \mathcal{C}} \psi_2(M_2 \beta_{sr}) \\
& = \sum_{<s,r> \in \mathcal{C}} \left[\beta_{sr} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 + \psi_2(M_2 \beta_{sr}) \right] \\
& = \mathcal{H}_2(\mathbf{0}, \beta; \boldsymbol{w}),
\end{aligned}
$$

one gets the following reduced prior energy:

$$
\boxed{\mathcal{H}_2^l(\mathsf{d}\boldsymbol{w}^l, \beta; \boldsymbol{w}) = \mathcal{H}_2(\mathbf{0}, \beta; \boldsymbol{w}) + M_2 \sum_{<n,m> \in \mathcal{C}^l} \left[\|\mathsf{d}\boldsymbol{w}_n^l - \mathsf{d}\boldsymbol{w}_m^l\|^2 B_{nm}^l + 2(\mathsf{d}\boldsymbol{w}_n^l - \mathsf{d}\boldsymbol{w}_m^l).\boldsymbol{W}_{nm}^l \right],} \tag{29}
$$

with:

$$
B_{nm}^l \triangleq \sum_{<s,r> \in \mathcal{C}_{nm}^l} \beta_{sr}, \quad \boldsymbol{W}_{nm}^l \triangleq \sum_{<s,r> \in \mathcal{C}_{nm}^l} \beta_{sr} (\boldsymbol{w}_s - \boldsymbol{w}_r).
$$

4.2 Energy minimization

The synopsis of the overall multigrid deterministic optimization at a given resolution level is the following:

Algorithm MIN$(\mathcal{H}, \boldsymbol{w}, f)$

- $\mathsf{d}\boldsymbol{w}^L \leftarrow 0$; ▷ *initialization*
- **for** $l = L$ to 0 ▷ *from coarsest grid to finest one*
  - **repeat**
    - $\delta \leftarrow \arg\min_\delta \mathcal{H}_1^l(\mathsf{d}\boldsymbol{w}^l, \delta; f, \boldsymbol{w})$; ▷ *computation of optimal weights δ (Eq. 30*
    - $\beta \leftarrow \arg\min_\beta \mathcal{H}_2^l(\mathsf{d}\boldsymbol{w}^l, \beta; \boldsymbol{w})$; ▷ *computation of optimal weights β (Eq. 31)*
    - **repeat**
      - $\mathsf{d}\boldsymbol{w}^l \leftarrow \text{WLS}(\mathsf{d}\boldsymbol{w}^l; \delta, \beta, f, \boldsymbol{w})$; ▷ *Weighted Least Squares minimization (Eq. 32)*
    - **until** convergence
  - **until** convergence
  - $\mathsf{d}\boldsymbol{w}^{l-1} \leftarrow [\Phi^{l-1}]^{-1} \circ \Phi^l(\mathsf{d}\boldsymbol{w}^l)$; ▷ *transmission of estimate to the next grid level*
- **end for**
- **return**$(\mathsf{d}\boldsymbol{w}^0, \delta, \beta)$;

The current reduced increment estimate $\mathsf{d}\boldsymbol{w}^l$ being fixed, we know that the optimal weight values are directly accessible. According to (15) in combination with energy definitions (9)(10), these values are:

$$
\hat{\delta}_s = \frac{1}{M_1} \phi_1' \left[\left(\boldsymbol{\nabla} f(s + dt\boldsymbol{w}_s, t + dt).\mathsf{d}\boldsymbol{w}_n^l + f_t(s, t, \boldsymbol{w}_s) \right)^2 \right], \ \forall s \in \mathcal{B}_n^l, \tag{30}
$$

$$\hat{\beta}_{sr} = \begin{cases} \dfrac{1}{M_2}\phi_2'\left[\|(\boldsymbol{w}_s + \mathsf{d}\boldsymbol{w}_n^l) - (\boldsymbol{w}_r + \mathsf{d}\boldsymbol{w}_m^l)\|^2\right] & \forall\ <s,r>\in \mathcal{C}_{nm}^l, \\ \dfrac{1}{M_2}\phi_2'\left[\|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2\right] & \forall\ <s,r>\in \mathcal{C}_n^l. \end{cases} \tag{31}$$

According to (31), the discontinuity variables β_{sr} located in between two neighboring blocks of $\mathcal{B}^l$ (i.e., $<s,r>\in \mathcal{C}_{nm}^l$ for some $<n,m>\in \mathcal{C}^l$) are the only ones to be iteratively updated as $\mathsf{d}\boldsymbol{w}^l$ evolves. The others only depend on $\boldsymbol{w}$ which is fixed along the whole multigrid procedure. Therefore, they can be computed right away at the first iteration of the current level. As soon as the values of auxiliary variables are computed and frozen, the energy function $\mathcal{H}^l(\mathsf{d}\boldsymbol{w}^l, \delta, \beta; f, \boldsymbol{w})$ is quadratic with respect to $\mathsf{d}\boldsymbol{w}^l$. Its minimization is equivalent to the resolution of a linear system whose solution is searched with an iterative Gauss-Seidel scheme similar to (5). All sites of S^l are repeatedly visited. If n is the current site of S^l, the reduced increment vector $\mathsf{d}\boldsymbol{w}_n^l = [\mathsf{d}u_n^l, \mathsf{d}v_n^l]^T$ is updated as follows:

$$\begin{cases} \mathsf{d}u_n^l \leftarrow \overline{u}_n^l - \dfrac{\gamma\langle \boldsymbol{F}_x^l | F_{xy}^l \overline{\boldsymbol{w}}_n + \boldsymbol{F}_t^l\rangle_n + A\overline{u}_n + \|\boldsymbol{F}_y^l\|_n^2\langle \boldsymbol{F}_x^l | \boldsymbol{F}_t^l\rangle_n - \langle \boldsymbol{F}_y^l | \boldsymbol{F}_t^l\rangle_n\langle \boldsymbol{F}_x^l | \boldsymbol{F}_y^l\rangle_n}{\gamma\left(\gamma + \|\boldsymbol{F}_x^l\|_n^2 + \|\boldsymbol{F}_y^l\|_n^2\right) + A} \\ \mathsf{d}v_n^l \leftarrow \overline{v}_n^l - \dfrac{\gamma\langle \boldsymbol{F}_y^l | F_{xy}^l \overline{\boldsymbol{w}}_n + \boldsymbol{F}_t^l\rangle_n + A\overline{v}_n + \|\boldsymbol{F}_x^l\|_n^2\langle \boldsymbol{F}_y^l | \boldsymbol{F}_t^l\rangle_n - \langle \boldsymbol{F}_x^l | \boldsymbol{F}_t^l\rangle_n\langle \boldsymbol{F}_x^l | \boldsymbol{F}_y^l\rangle_n}{\gamma\left(\gamma + \|\boldsymbol{F}_x^l\|_n^2 + \|\boldsymbol{F}_y^l\|_n^2\right) + A} \end{cases} \tag{32}$$

with

$$\overline{\boldsymbol{w}}_n^l \triangleq \frac{\sum_{m\in\nu(n)} B_{nm}^l \mathsf{d}\boldsymbol{w}_m^l - \boldsymbol{W}_{nm}^l}{\sum_{m\in\nu(n)} B_{nm}^l} \triangleq [\overline{u}_n^l, \overline{v}_n^l]^T,$$

$$\gamma \triangleq \frac{\alpha M_2}{M_1} \sum_{m\in\nu(n)} B_{nm}^l,$$

$$A \triangleq \|\boldsymbol{F}_x^l\|_n^2 \|\boldsymbol{F}_y^l\|_n^2 - \langle \boldsymbol{F}_x^l | \boldsymbol{F}_y^l\rangle_n^2.$$

Note that in the above expressions, $\boldsymbol{F}_x^l$, $\boldsymbol{F}_y^l$, $\boldsymbol{F}_t^l$ vectors, and F_{xy}^l matrices as well, are displayed without $(n, \boldsymbol{w})$ for the sake of concision.

4.3 Monoresolution case

It is quite interesting deriving the multigrid smoothing energy term in case $\boldsymbol{w} \equiv \boldsymbol{0}$. This occurs at the coarsest resolution level. This is of particular interest in case of small displacements, when there is no need of incremental estimation (i.e., a *monoresolution* setup is sufficient).

From 28, it comes in this case:

$$\mathcal{H}_2^l(\mathsf{d}\boldsymbol{w}^l, \beta; \boldsymbol{0}) = M_2 \sum_{<n,m>\in\mathcal{C}^l} \sum_{<s,r>\in\mathcal{C}_{nm}^l} \beta_{sr}\|\mathsf{d}\boldsymbol{w}_n^l - \mathsf{d}\boldsymbol{w}_m^l\|^2 + \sum_{<s,r>\in\mathcal{C}} \psi_2(M_2\beta_{sr}).$$

As seen around (31), the optimal values of the inner-block auxiliary variables depend only on $\boldsymbol{w}$ which is zero. Consequently, they can be set at once at their optimal values which is 1. For

the same reason, the optimal value of the auxiliary variable attached to a neighboring site pair straddling blocks $\mathcal{B}_n^l$ and $\mathcal{B}_m^l$ depends only on $\mathsf{d}\boldsymbol{w}_n^l$ and $\mathsf{d}\boldsymbol{w}_m^l$. Therefore, only one auxiliary variable, denoted β_{nm}^l has to be used per neighboring site pair $<n,m>$ of the reduced grid. The relevant prior turns out to be:

$$\mathcal{H}_2^l(\mathsf{d}\boldsymbol{w}^l,\beta^l) = \sum_{<n,m>\in\mathcal{C}^l} |\mathcal{C}_{mn}^l| \left[M_2\beta_{nm}^l \|\mathsf{d}\boldsymbol{w}_n^l - \mathsf{d}\boldsymbol{w}_m^l\|^2 + \psi_2(M_2\beta_{nm}^l)\right] \ ,$$

where $\beta^l \triangleq \{\beta_{nm}^l, <n,m>\in \mathcal{C}^l\}$ is a reduced set of auxiliary variables lying on the dual lattice of S^l. By minimizing w.r.t. β^l, one gets the underlying robust prior energy:

$$H_2^l(\mathsf{d}\boldsymbol{w}^l) = \sum_{<n,m>\in\mathcal{C}^l} |\mathcal{C}_{mn}^l|\, \rho_2(\|\mathsf{d}\boldsymbol{w}_n^l - \mathsf{d}\boldsymbol{w}_m^l\|).$$

It is striking that we thus derive a prior energy at grid level l which exhibits exactly the same form as the original one (10); including the neighborhood system: if S is equipped with a first-order (resp. second-order neighborhood) system as for the definition of the smoothing prior, one ends up with the same neighborhood system on reduced grid S^l. The only difference comes from appearing weighting parameters. The smoothing energy "contribution" of pair $<n,m>\in\mathcal{C}^l$ is multiplied by $|\mathcal{C}_{mn}^l|$, i.e., the number of neighboring site pairs of S astride blocks $\mathcal{B}_n^l$ and $\mathcal{B}_m^l$. Here, we chose $2^l \times 2^l$ blocks, S being equipped with the first-order neighborhood system. It then comes $|\mathcal{C}_{mn}^l| = 2^l$. In case of second-order neighborhood system, this weight would have grown to $(3 \times 2^l - 2)$ for "horizontal" and "vertical" block pairs, whereas it would have remained unity for "diagonal" pairs of blocks (which are connected through a *single* diagonal pair of neighboring sites.

We have completely described a multigrid reweighted least squares minimization algorithm which is here devoted to robust optical flow estimation. The use of such a multigrid iterative relaxation in the present context enables us to build an optical flow estimator which is formally speaking similar to Horn and Shunck's estimator. In contrast, it is robust to the failures of the OFCE-based model, and it is able to take into account (localization and preservation) the discontinuities of the optical flow.

5 Coupling with an object-oriented motion segmentation

In this section we briefly introduce an extension of the model to couple the motion estimation process with an "object-oriented segmentation". By this, we mean that we aim at simultaneously extracting the silhouette of one dynamical entity [5]. This is of great interest for

[5]The extension to a *fixed* number of such entities would be straightforward. The estimation of an *unknown* and *varying* number of regions is a though problem of its own. We do not address here this issue.

interpreting, structuring and better estimating the optical flow information [11, 43]. Hereafter, the segmentation we consider is defined as a *single closed curve* lying in the image plane, which does not self-intersect. Some discretization scheme allows to associate with it a unique partition of the pixel set in terms of "interior", R, and "exterior", $\bar{R} = S - R$. In the following R will be called *a segment*. The precise definition and parametrization of this curve may be of different natures, depending on the problem at hand, on the amount of prior knowledge which is available or wanted to be used, and on the affordable computational burden. The extraction of moving objects of a precise nature (e.g., vehicles in traffic sequences), and the extraction of the region exhibiting the largest motion in image sequences with various contents, should be approached differently from this point of view.

A parametrization of low dimension can be used to specify a strong geometric prior knowledge of the region shape. This is allowed by the *deformable template* framework proposed by Grenander [21, 22, 23]. Conversely, a "weak" regularizing prior can be captured with plane curves controlled by a large number of parameters. This situation is thoroughly addressed within the framework of *active contour models* and *snakes* [14, 29]. Note the two important classes of approaches that we have just evoked are eventually quite overlapping.

Originally used in still image segmentation, these closed curve models appear as powerful tools to cope with motion segmentation. Recently, different extensions of deformable curve formalism to motion detection and tracking have been described [3, 4, 8, 17, 30, 32, 35, 38, 42].

Our purpose here is neither to choose among the different deformable shape models, nor to propose a new one. We would rather like to show how such a curve-based segmentation may be connected to the optical flow estimator we propose. Therefore, in the coming section, we remain general by neither restricting ourselves to a specific family of possible curves, nor to a specific prior knowledge.

As for the experiments, we will demonstrate the feasibility and the interest of the approach by using two simple types of object-oriented segmentation models which illustrate two extreme cases of parameterizations.

5.1 Energy design

The extension of the energy-based estimation model is obtained by adding two terms to the global energy function $\mathcal{H}$ [6]. The first one, E_{prior}, captures the *a priori* knowledge about the *segmenting curve*. As we said, it can be of very different natures. The second one, $E_{interact}$, specifies the mode of interaction between the segment and the rest of the estimation model (i.e., velocity field, auxiliary variables and data). Different ways of interaction may be considered:

[6] Even though the superscript remains omitted, we still suppose in the coming developments that some resolution level k is concerned

- The segment interacts directly with $\boldsymbol{w}$ by "cutting" the regularization through its border. In this case it acts as a set of binary line processes similar to those in [19, 24], which would be constrained to lye along a closed curve;

- The segment interacts indirectly with the velocity field through the auxiliary variables β and/or δ. An interaction with β would favor velocity discontinuities along the border of the moving region, and conversely, it would drive the curve toward the already computed discontinuities of the current optical flow estimate.

- The segment interact with the data lying along its border and/or within the span of its interior. This could allow to capture the fact that the boundary of a moving region is very likely to exhibit large photometric discontinuities, while the interior should exhibit, to some extent, a noticeable spatial/temporal brightness homogeneity.

Each of these possibilities has some appeal, and they could possibly be fruitfully combined. For instance, the curve could simultaneously be driven by the data toward photometric discontinuities while interacting directly or indirectly with the regularization of the velocity field.

Our purpose here is to introduce a first attempt to equip our complete estimation model with an object-oriented interacting segmentation device: we have chosen so far an interaction mode which leads to a simple energy formulation, and to easy computations. The segment will only interact with the estimation process through the β's auxiliary variables. The corresponding cost function component $E_{interact}(R,\beta)$ exhibits two terms: the first one is proportional to the mean value of β_{sr}s over $\mathcal{C}_{\partial R} \triangleq \{<s,r> \in R \times \bar{R}\}$. It then (a) favors low values (close to zero) of discontinuity auxiliary variables along the border of the segment, and (b) drives the curve toward the more significant gathering of low-valued β_{sr}. The second one is proportional to the opposite mean value of β_{sr}s over $\mathcal{C}_R \triangleq \{<s,r> \subset R\}$. Its role is (a) to favor large values (close to unity) inside and (b) to make the curve surround areas with uniform velocity. The global energy of the extended model is designed as follows:

$$\mathbb{H}(\mathrm{d}\boldsymbol{w}, \delta, \beta, R; f, \boldsymbol{w}) \triangleq \mathcal{H}(\mathrm{d}\boldsymbol{w}, \delta, \beta; f, \boldsymbol{w}) + E_{prior}(R) + E_{interact}(R,\beta), \tag{33}$$

with:

$$E_{interact}(R,\beta) = \frac{\mu_1}{|\mathcal{C}_{\partial R}|} \sum_{<s,r> \in R \times \bar{R}} \beta_{sr} - \frac{\mu_2}{|\mathcal{C}_R|} \sum_{<s,r> \subset R} \beta_{sr}, \tag{34}$$

with some positive parameters μ_1 and μ_2.

As for the class of admissible segments, and associated prior energy, we have considered two extreme cases in our experiments:

- A case of tight geometric constraint associated with a weak prior on the shape: the segmenting curve is a convex quadrilateral whose prior energy is proportional to $\frac{\text{perimeter}^2}{\text{surface}}$, hence favoring "compact" shapes.

- A case of lose geometric constraint associated with a classical Minimum Length Description prior [31, 36]: the segmenting curve is any closed non self-intersecting curve whose cost is proportional to its length (above a certain threshold [7]).

5.2 Energy minimization

The alternate minimization spirit is still considered. Compared to the segmentation-free optical flow estimation of §4.2, two extra features have to be described: (1) the new computation rules of the optimal discontinuity auxiliary variables (since they now interact not only with the velocity field as before, but also with the segmenting curve); (2) the computation of this segmenting curve with respect to the current optical flow under estimation.

For a given velocity field $\boldsymbol{w}+\mathsf{d}\boldsymbol{w}$, and a given segment R, the optimal discontinuity weights are not any more given by (31). For each neighbor pair $<s,r>$, the actual problem to solve is the following minimization:

$$\min_{\beta_{sr}\in(0,1]} \mathbb{H}(\mathsf{d}\boldsymbol{w},\delta,\beta,R;f,\boldsymbol{w}) = \min_{\beta_{sr}\in(0,1]} \left[\beta_{sr}\left(M_2\|\Delta\boldsymbol{w}_{sr}\|^2+\frac{\mu_1'}{|\mathcal{C}_{\partial R}|}\mathbf{1}_{\partial R}(s,r)-\frac{\mu_2'}{|\mathcal{C}_R|}\mathbf{1}_R(s,r)\right)+\psi_2(M_2\beta_{sr})\right], \quad (35)$$

where, $\Delta\boldsymbol{w}_{sr}=(\boldsymbol{w}_s+\mathsf{d}\boldsymbol{w}_s)-(\boldsymbol{w}_r+\mathsf{d}\boldsymbol{w}_r)$ as a notational convenience, $\mathbf{1}_R(s,r)$ (resp. $\mathbf{1}_{\partial R}(s,r)$) equals 1 if $<s,r>\subset R$ (resp. $<s,r>\in R\times\bar{R}$), and 0 otherwise, and $\mu_i'\triangleq\frac{\mu_i}{\alpha}$, $i=1,2$. One gets the new optimal weight computation rule (see Appendix B for the proof):

$$\text{if } <s,r>\subset\bar{R},\quad \hat{\beta}_{sr}=\frac{1}{M_2}\phi_2'\left[\|\Delta\boldsymbol{w}_{sr}\|^2\right]; \quad (36)$$

$$\text{if } <s,r>\in R\times\bar{R},\quad \hat{\beta}_{sr}=\frac{1}{M_2}\phi_2'\left[\|\Delta\boldsymbol{w}_{sr}\|^2+\frac{\mu_1'}{M_2|\mathcal{C}_{\partial R}|}\right]; \quad (37)$$

$$\text{if } <s,r>\subset R\ ,\hat{\beta}_{sr}=\frac{1}{M_2}\phi_2'\left\{\left[\|\Delta\boldsymbol{w}_{sr}\|^2-\frac{\mu_2'}{M_2|\mathcal{C}_R|}\right]^+\right\}, \quad (38)$$

where $[.]^+\triangleq\max\{0,.\}$. Basically, the optimal weight computation rule is only changed along and within the moving region, according to a simple shift of the argument of ϕ_2' (provided this argument remains positive. See (38).). Along the border this (positive) shift results in a decrease of the weights (i.e., smoothing reduction) which is all the more important that the border is short. Inside the region, the (bounded negative) shift results in an increase of the weights (i.e., smoothing accentuation), with a saturation at 1. In this context, Leclerc's estimator yields a very convenient updating rule. As a matter of fact, since $\frac{1}{M_2}\phi_2'(v)=\exp(-M_2v)$, the computation of the optimal weight turns out to be simply related to the segmentation-free optimal weight ($\hat{\beta}^0_{sr}\triangleq\frac{1}{M_2}\phi_2'\left[\|\Delta\boldsymbol{w}_{sr}\|^2\right]$) through a single multiplication (see Fig. 5.2 for a plot

[7] This threshold is necessary not to strongly favor curve with length close to zero

example):

$$\text{if } <s,r> \subset \bar{R},\ \hat{\beta}_{sr} = \hat{\beta}^0_{sr}; \tag{39}$$

$$\text{if } <s,r> \in R \times \bar{R},\ \hat{\beta}_{sr} = \underbrace{\exp(-\frac{\mu'_1}{|\mathcal{C}_{\partial R}|})}_{\triangleq K_{\partial R}<1} \hat{\beta}^0_{sr}; \tag{40}$$

$$\text{if } <s,r> \subset R, \hat{\beta}_{sr} = \min\{1, \underbrace{\exp(\frac{\mu'_2}{|\mathcal{C}_R|})}_{\triangleq K_R>1} \hat{\beta}^0_{sr}\}. \tag{41}$$

A multigrid version, with velocity field increments being piecewise constant over blocks, is easy to derive. It will not be detailed here.

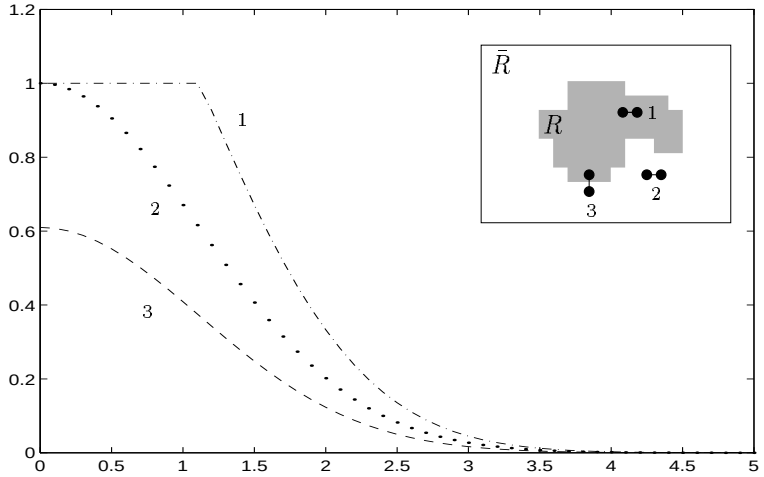

Figure 7: Example of modified interaction functions on the basis of Leclerc estimator ($\sigma = 2.5$). For $<s,r> \subset \bar{R}$ (resp. $<s,r> \in R \times \bar{R}$), we took $\frac{\mu'_2}{|\mathcal{C}_R|} = 0.5 \Rightarrow K_R = 1.65$ (resp. $\frac{\mu'_1}{|\mathcal{C}_{\partial R}|} = 0.5 \Rightarrow K_{\partial R} = 0.61$)

The minimization of the energy according to the segment is addressed in two successive ways. First, on the coarsest grid ($l = L$) of the coarsest resolution ($k = N$) where the complete procedure starts and where the dimensionality of the problem is drastically reduced, one can sweep efficiently over the set of possible segmentations. A crude estimate of the segment is thus obtained at low cost by a stochastic algorithm, with *no need of any manual initialization.* The location and shape of the segment is then refined deterministically through the following grids of resolution N, and through the finest grids ($l = 0$) of the following resolution levels $N-1, \ldots, 0$. For both kinds of segment updating, the velocity vectors being fixed, an iterative scheme is used which considers at each step different possible segments R *along with the associated optimal weights* $\hat{\beta}$ provided by (36),(37) and (38). The part of energy actually concerned by the updating

process reduces to the one involving the segment and/or the discontinuity auxiliary variables:

$$E(R) \triangleq E_{prior}(R) + \sum_{<s,r>\in\mathcal{C}} \left[\hat{\beta}_{sr} \left(M_2 \|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{|\mathcal{C}_{\partial R}|} \mathbf{1}_{\partial R}(s,r) - \frac{\mu_2'}{|\mathcal{C}_R|} \mathbf{1}_R(s,r) \right) + \psi_2(M_2 \hat{\beta}_{sr}) \right].$$

In the first stage ($l = L, k = N$), in absence of any "initial guess", the set of possible segmentations has to be explored thoroughly. This is done by using a simulated annealing algorithm based on Metropolis dynamics. Given the current segment R at step n, a new segment R' (with optimal weights β') is proposed by randomly applying a global deformation (rotation, translation and scaling) to R [22, 30, 38]. The new segment is accepted with probability $\exp\{-[E(R') - E(R)]^+ / T_n\}$, according to a geometric cooling $T_n = T_0 \times (0.99^n)$.

For the second type of updating, an initial guess is always available by properly interpolating the segment obtained at the previous grid level (if $k = N$) or resolution (if $k < N$). Given this initial guess, we simply seek a *local minimizer* of the energy nearby this guess, proceeding with "small" *local deformations*. At each step, a set of deformations applied to the same section of the curve are considered. The one which provides the largest energy decrease is chosen. In the case of quadrilateral segments for instance, one of the four vertices is picked and displacements in a 3×3 window around its current location are considered, and compared in terms of global energy. The sketch of the overall method's synopsis is presented in Fig. 8.

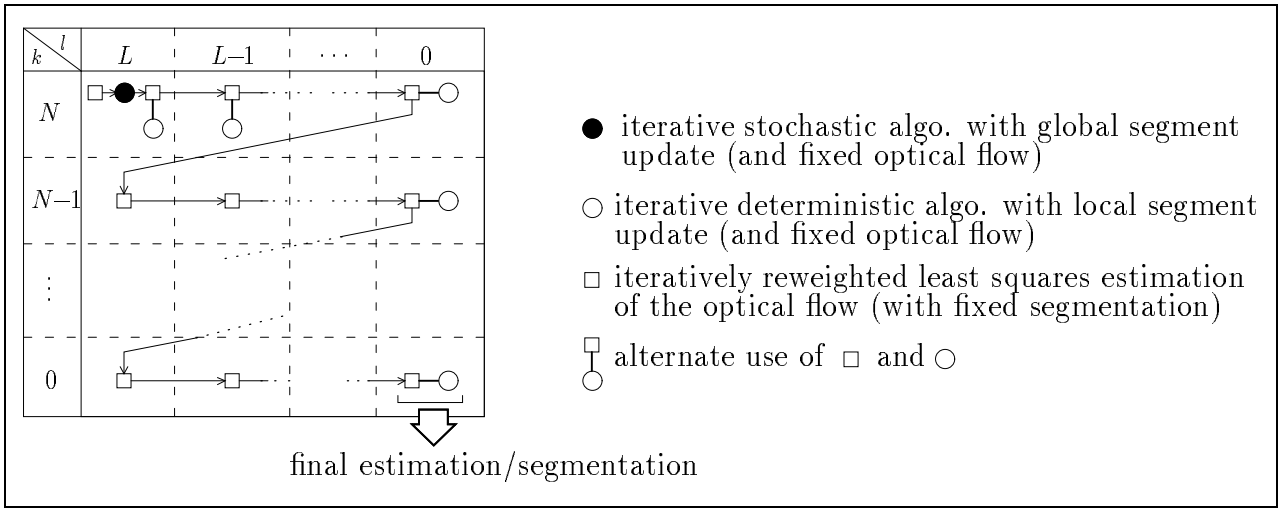

Figure 8: Schematic synopsis of the complete estimation/segmentation method

Before turning to experimental results, let us come back to the energy change computation. In case of Leclerc's estimator, using a first order expansion for inner contour sites such that $\|\Delta \boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{M_2|\mathcal{C}_R|} < 0$, one gets the following simple term to compute:

$$\tilde{E}(R) \triangleq E_{prior}(R) + (1 - K_{\partial R}) \sum_{<s,r>\in R\times\bar{R}} \hat{\beta}^0_{sr} + (1 - K_R) \sum_{<s,r>\subset R} \hat{\beta}^0_{sr},$$

for any segment to compare (see Appendix C for details). Then the update is ruled by energy variations $E(R') - E(R) \approx \tilde{E}(R') - \tilde{E}(R)$ which are very easy to compute.

	Yosemite	Parking lot	Calendar
number of resolution levels $(N+1)$	2	1	3
number of grid levels at each resolution $(L+1)$	4	5	5
smoothing parameter α	320	200	200
parameters σ_1 tuning ρ_1	6	6	7
parameters σ_2 tuning ρ_2	0.7	0.2	0.4

Table 1: Parameter values for experiments on the three sequences

6 Experimental results

In this section we present results of optical flow estimation alone (§6.1), and results of joint estimation/segmentation (§6.2).

6.1 Optical flow estimation

The optical flow estimation model presented in the first part of this paper has been validated both on synthetic and real-world sequences. The first one, Yosemite (Fig. 9a-b), is the most complex (though synthetic) sequence from the comparative study by Barron *et al.* [2] for which a "ground-truth" exists. The two other sequences are real. They exhibit far more motion discontinuities and OFCE violations. Therefore, they are probably more adapted than Yosemite to demonstrate the nice characteristics and abilities of our method, except that there is no ground-truth for them. Figure 12a shows a Parking lot sequence which involves two cars moving in the foreground. The camera pans the scene as the wind shakes the trees in the background. Calendar (Fig. 13a) is a TV sequence involving large displacements. It includes several different moving objects and a horizontal panning camera motion. The calendar translates vertically and the toy-train pushes a rolling ball.

The choice of the two robust estimators ρ_1 and ρ_2 has been based on heuristic considerations arising from our experience. Since frequent and large deviations from the brightness constancy assumption (which supports the OFCE) are more than likely to occur, a strongly saturating estimator seems to be well suited to the corresponding component of the energy function. That's why we selected Leclerc's estimator [31] (Fig. 4). On the contrary, a softer saturation seems to be better as far as regularization is concerned. The extreme (and degenerated) case of the truncated quadratic as a hardly saturating cost function, introduces bad local minima, resulting in a painful oscillatory behavior of the alternate minimization procedure. This still holds to some extent for robust estimators close to this degenerated figure. Therefore, we chose Geman and McClure's estimator [20] to be embedded within the smoothness constraint due to the slower decreasing rate of its derivative function (Fig. 4).

The values of the different parameters for the three sequences are gathered in table 1. Note that for Parking lot, the displacements are small enough not to use the incremental multiresolu-

tion setup ($N = 0$). For the other sequences, the same values of parameters $\alpha, \sigma_1, \sigma_2$ have been selected throughout the multiresolution structure. As for the parameter tuning throughout the multigrid structure associated to a given resolution level, it is *implicitly and automatically* done within the energy derivations presented in §4.1.

Following [2], quantitative comparative results on Yosemite are provided for different algorithms. For each estimate, the deviation with respect to the "real" flow is measured at each pixel location by converting the 2D vectors into 3D unit vectors, and by computing the angle between them:

$$\text{error}(s) \triangleq \arccos\left(\frac{u_s.u_s^{real} + v_s.v_s^{real} + 1}{\sqrt{\|\boldsymbol{w}_s\|^2 + 1}\ \sqrt{\|\boldsymbol{w}_s^{real}\|^2 + 1}}\right).$$

For some methods reported in [2], estimates are only available at so called "reliable" locations. The percentage of such locations is the "density" of the estimate. The errors are actually computed only at these locations. Table 2 lists the average and standard deviation of these angular discrepancies for different algorithms. The 5 top-lines recall results presented by Barron *et al.* They concern two different versions of Horn and Schunck's algorithms (whose model is the basis of our approach), the best full-density algorithm (Uras *et al.*) and the two algorithms yielding the best results, but with reduced densities (Lucas and Kanade, Fleet and Jepson). Other authors have provided similar comparisons, but on a sub-sequence where the sky was removed. As a matter of fact this region is extremely tricky due to the complex luminance and shape evolution of the moving clouds before the sun. The second part of the table compares our method to those by Szeliski and Coughlan [44] and by Black [6] on this reduced sequence.

Technique	Average error	Standard deviation	Density	
Horn and Schunck (original)	31.69^o	31.18^o	100%	
Horn and Schunck (modified)	9.78^o	16.19^o	100%	
Uras *et al*	8.94^o	15.61^o	100%	
Lucas and Kanade	4.28^o	11.41^o	35.1%	
Fleet and Jepson	4.63^o	13.42^o	34.1%	
Robust multigrid	$\mathbf{5.37^o}$	$\mathbf{8.19^o}$	**100%**	
Szeliski and Coughlan	2.45^o	3.05^o	100%	} without sky
Black	3.52^o	3.25^o	100%	} without sky
Robust multigrid	$\mathbf{2.34^o}$	$\mathbf{1.45^o}$	**100%**	} without sky

Table 2: Comparative results on Yosemite

On the complete scene, our method provides a *dense* estimate almost as good as those obtained with the best (non-dense) mentioned methods. In addition, the obtained standard deviation is the smallest one. On the sub-scene, the average error is slightly lower than the one obtained by Szeliski and Coughlan, with a standard deviation reduced of one half. The difference between our estimate and the "real" flow is displayed in Fig. 9c (subsampled and

multiplied by 10). It clearly appears that most of the discrepancies are gathered around the two moving clearings in the clouds, which the sunlight breaks through.

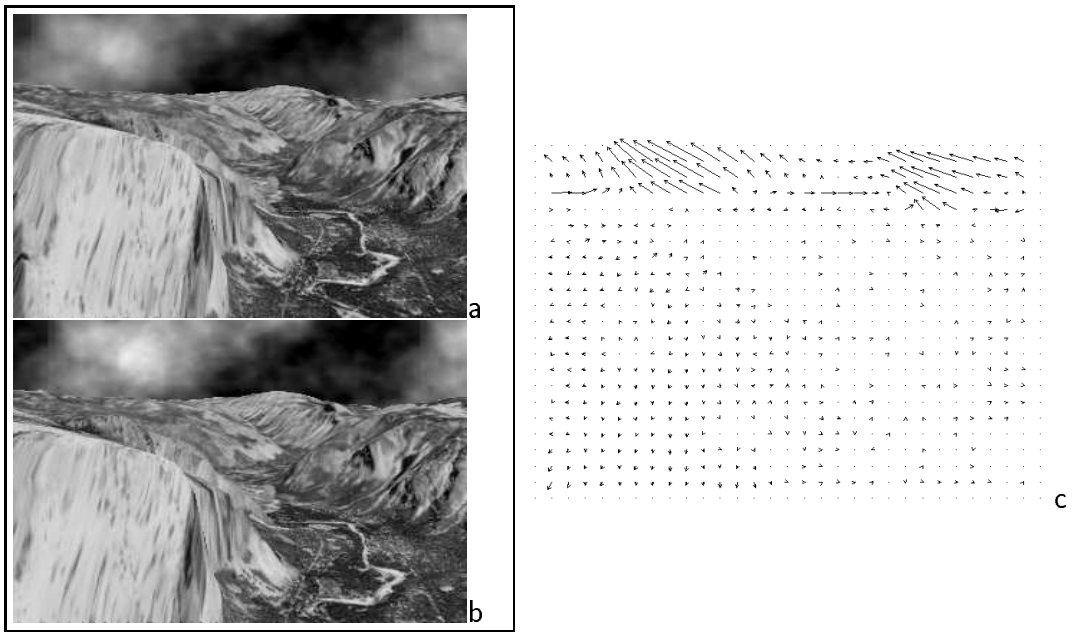

Figure 9: (a-b) Two frames of Yosemite (c) difference ($\times 10$) with the "true" flow.

To demonstrate the low sensitiveness of our method to parameters, we run it on Yosemite for 1000 triples $(\alpha, \sigma_1, \sigma_2) \in [140, 320] \times [3, 12] \times [0.1, 1]$. In Fig. 10, we plot the corresponding histograms of average angular errors for the complete scene and the sub-scene. In both cases, a high robustness is exhibited. This is also noticeable in the plot of the average error versus (σ_1, σ_2) for fixed $\alpha = 320$ (Fig. 11).

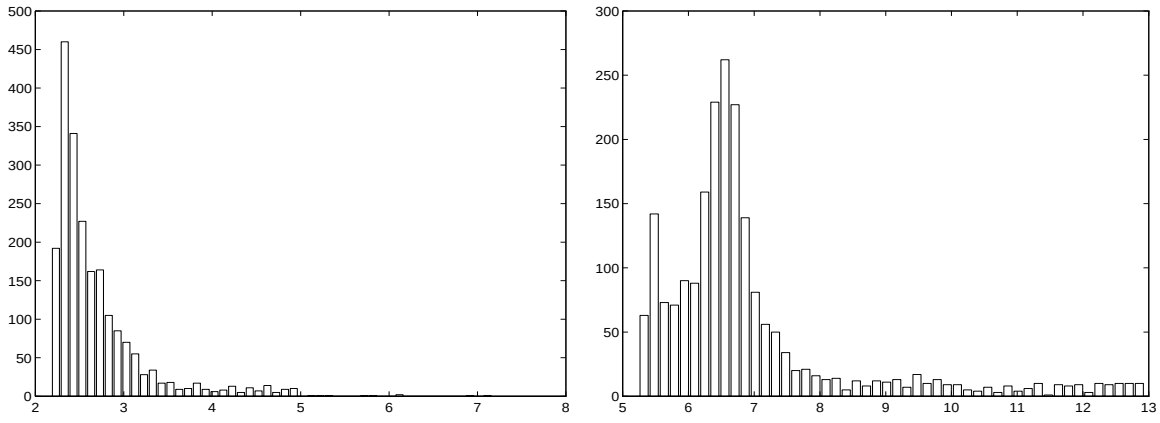

Figure 10: Average angular error histograms for 1000 parameter triples $(\alpha, \sigma_1, \sigma_2) \in [140, 320] \times [3, 12] \times [0.1, 1]$: without the sky (left) and with the sky (right)

Figures 12 and 13 present final flow estimates and discontinuity weights at different resolutions and/or grid levels for Parking lot and Calendar. The discontinuity weights are displayed on

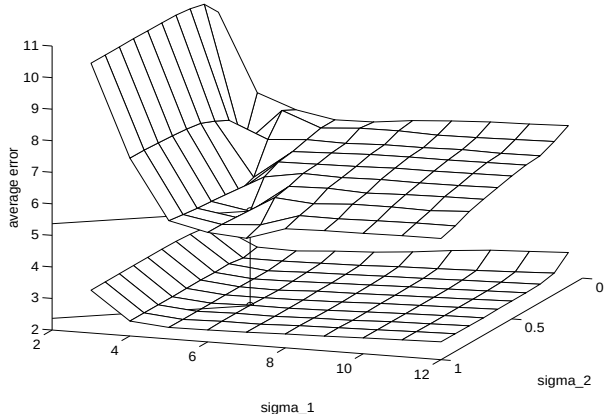

Figure 11: Average angular error for 100 parameter triples $(\sigma_1, \sigma_2) \times [3, 12] \times [0.1, 1]$, with $\alpha = 320$: without sky (lower surface) and with the sky (upper surface). In both cases, the minimal error is obtained for $\sigma_1 = 6$ and $\sigma_2 = 0.7$

255 grey levels on sites of the dual lattice: white points represent zero weights (i.e., maximal discontinuity) and black points are for unity weights (i.e., no discontinuity).

For all sequences we obtained good results. Visually, the recovered optical flow fields are close to the idea one can picture oneself of the "real" apparent motion field . The location of spatial discontinuities fits with a good accuracy the boundaries of the different motion regions (e.g., the first car in Parking lot, the train in Calendar). One can also notice that the motion estimation remains consistent in tough parts of the dynamic scenes at hand. For example, the motion of the calendar is quite well estimated despite highly textured portions exhibiting periodical patterns (like the drawing of houses). Another interesting example is found on the left part of the front car windshield in Parking lot. The motion of a patch of specular reflection does not disturb the car velocity estimation in the neighborhood. From the final velocity field estimate, it appears as an independent moving patch which goes in a direction opposite to the one of the car motion. This demonstrates how the estimator behaves well in cases the OFCE is well known to be non valid.

We believe that the good quality of these results is partly due to the multigrid structure of the estimator. It produces good intermediate coarse flows. Besides, it allows to extract "long structures" of discontinuities. This is particularly noticeable in Calendar (Fig. 13d-e).

Finally, note the whole multiresolution/multigrid algorithm converges quickly, since only ten or so low cost iterations are required at each grid level.

6.2 Coupling with the segmentation

We experimented the augmented model with a single segmenting closed curve on real traffic sequences involving clear dynamical entities (i.e., vehicles apparently moving). We present results on two sequences. The first one is Parking lot which has already been introduced. The

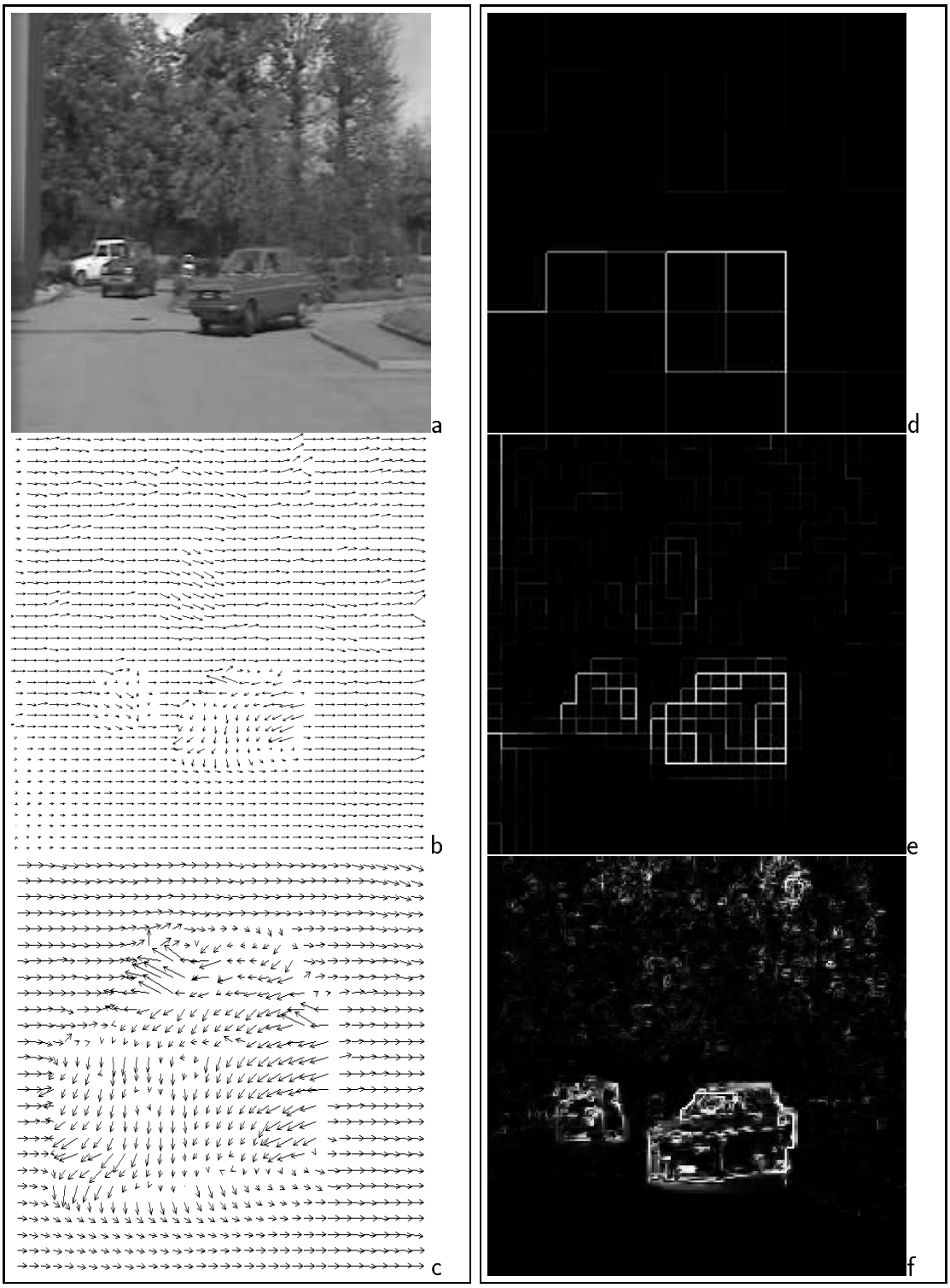

Figure 12: Results on Parking lot: (a) one frame (b) flow estimate (c) zoom on the foreground car (d-e-f) discontinuity weights at resolution level $k = 0$, on grid levels $l = 4, 2, 0$.

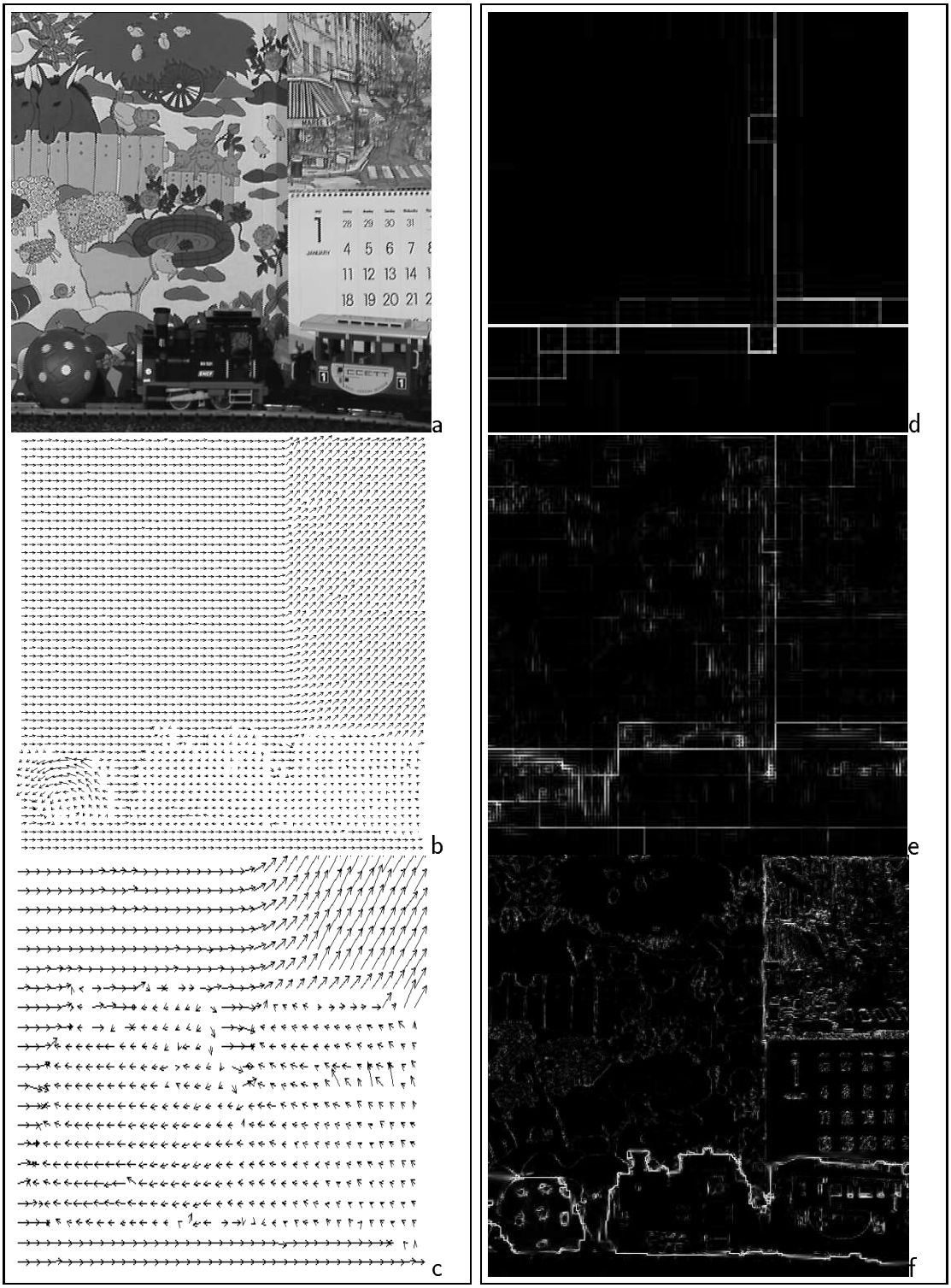

Figure 13: Results on Calendar: (a) one frame (b) flow estimate (c) zoom on the train (d-e-f) discontinuity weights in $\Omega^{2,3}$, $\Omega^{1,4}$ and Ω.

second one, called Street, is shot from a car driving in a stream of moving vehicles (Fig. 14). The motion of the camera results in a divergent motion component in the image. Within this field, the moving entity with the largest apparent motion corresponds to the car entering the image plane in the foreground, and moving away from the camera.

Figure 14: Two consecutive frames of Street sequence.

As for the choice of the two robust M-estimators, we explained in §6.1 why we chose Geman and McClures's function in the smoothing term. However, as pointed out in §5.2, Leclerc's estimator, due to its exponential nature, yields simpler computations as far as the smoothing energy term is concerned in the estimation/segmentation model. Therefore, in place of Geman and McClure's estimator, we made this convenient choice for ρ_2 in experiments about the augmented model.

We experimented the augmented model for a single segmentating curve. As explained in §5.1, two "extreme" cases of *a priori* on the curve have been considered: in the first case, the curve is constrained to be a convex quadrilateral (thus defined through eight parameters) along with a prior favoring compact shapes; in the second case, the curve is only constrained to be closed and non-self-intersecting, and a MDL-type prior is used to stay around a "reasonable" border length.

Values of the different parameters for the two sequences are gathered in table 3. For Street, the same values of parameters $\alpha, \sigma_1, \sigma_2, \mu_1'$ and μ_2' have been selected throughout the three level multiresolution structure.

Similar results of good quality are obtained with both kinds of curves. A low dimension parametrization (as with the quadrilaterals) yields computations of lower cost, while allowing to capture some strong geometric knowledge (if available) on the shapes of the moving entities. At the same time, this can become a drawback, if the prior knowledge is not sufficient or the restricted family of curves is not well suited to the scene content (compare c and f in Fig. 15 and Fig. 16). In our case, quadrilaterals are not able to completely fit the complex shape of moving vehicles. As a consequence, the "importance" of the border-based energy risks to

Figure 15: Single object motion segmentation on Parking lot: (a-b-c) with an *unconstrained closed curve* at grid levels 4,2 and 0 (d-e-f) with a *convex quadrilateral* at grid levels 4,2 and 0.

Figure 16: Single object motion segmentation on Street: (a-b-c) with an *unconstrained closed curve*, in $\Omega^{2,2}$, Ω^2 and Ω (d-e-f) with a *convex quadrilateral* in $\Omega^{2,2}$, Ω^2 and Ω.

	Parking lot	Street
number of resolution levels $(N+1)$	1	3
number of grid levels at each resolution $(L+1)$	5	5
smoothing parameter α	200	200
scale parameter σ_1	8	8
scale parameter σ_2 for the *quadrilateral*	0.1	0.1
scale parameter σ_2 for the *unconstrained curve*	0.3	0.3
segment's border parameter μ_1'	0.5	0.5
segment's interior parameter μ_2' for the *quadrilateral*	0.05	0.05
segment's interior parameter μ_2' for the *unconstrained curve*	0.5	0.5

Table 3: Parameter values for segmentation experiments

shrink: the region-based term has to be reduced by tuning μ_2' to a very low value (see Table 3) in order to keep a proper balance. Otherwise, the template would be mainly driven by region-wise flow uniformity, risking to get stuck around spurious locations such as parts of still background in case of non-moving camera. These problems can hopefully be circumvented by using less constrained curves, but at increased cost: the computation burden is usually larger and the stability of the iterative scheme can become a critical issue.

7 Conclusion

In this paper, we have presented a comprehensive multiresolution/multigrid framework for optical flow estimation and object-oriented motion-based segmentation.

The estimation problem is expressed as the global minimization of an energy function which involves robust estimators to avoid spatial over-smoothing and to attenuate the influence of large data model deviations. The minimization is processed through a multigrid algorithm which consists in imposing successively weaker and weaker constraints on the searched estimates. Applied to a dual formulation of the original energy function, this method leads to a multigrid iteratively reweighted least squares minimization which is efficient in term of convergence rate and in term of quality of the produced estimates. It is worth noting that this is a *general purpose* multigrid approach which can be easily applied to most of image analysis objective functions which involve robust M-estimators. Besides, it allows to define efficient and original *parallel* relaxation algorithms [34].

In the motion estimation context, one of the nicest features of the approach is that it gives access to a consistent and *meaningful* sequence of finer and finer configurations at each single level of data resolution. The coarser grain estimates, which are far easier to compute due to their reduced dimensionality, reveal large discontinuity structures of the apparent motion field. This kind of significant reduced estimates cannot be produced by continuation methods

since the earliest estimates they provide are, to some extent, smoothed versions of their final estimates [5, 9, 31].

We take advantage of these compact and structured information by introducing a closed curve-based device, which allows to recover with improved accuracy the location of spatial discontinuities and to naturally handle *edge grouping* to get some *object-oriented motion segmentation*. The proposed model can support any kind of parameterized or non-parameterized family of curves, equipped with any prior energy function.

We thus contribute here to the efforts to define global approaches for a joint and cooperative handling of two important interleaved issues of motion analysis. As far as our object-oriented motion segmentation is concerned, it is worth noting that neither knowing the motion of the camera (if any), nor initializing manually the curve are required.

Among the different possible ways of interaction between the segmenting curve and the other variables (velocities, weights, data), we chose one of the simplest in which the curve only interacts with the discontinuity auxiliary variables. Further extensions of this work will aim at driving in addition the segmenting curve by the data through:

- The spatial and temporal variations of luminance around and inside the segment, as proposed in [38] for motion detection with still camera;
- The data outlier auxiliary variables whose low values are usually concentrated in *occlusion areas* which surround moving entities.

The interaction mechanism we designed with segmenting closed curve, also suggests to use it with *open* curves. In this case, one would not segment the images, but this would be useful for preserving and refining through grid and resolution levels, the precious information captured by the (independent) discontinuity weights on smallest grids.

A Dual formulation of energy-based robust estimation

Let $\phi(v) \triangleq \rho(\sqrt{v})$ be continuously defined on $\mathbb{R}^+$. It is continuously differentiable on $\mathbb{R}^{+*}$ and strictly concave. Beside, $\lim_{v\to+\infty} \phi'(v) = \lim_{v\to+\infty} \frac{\rho'(\sqrt{v})}{2\sqrt{v}} = 0$ since $\lim_{u\to+\infty} \rho'(u) < \infty$. Hence $\phi'(v)$ is a decreasing bijection from $\mathbb{R}^+$ into $(0, M]$, with $M \triangleq \lim_{u\to 0^+} \frac{\rho'(u)}{2u}$ (not necessary finite).

For any $v \in \mathbb{R}^+$, the tangent to the graph of ϕ at point $(v, \phi(v))$ has the equation $y = \phi'(v)x + [\phi(v) - v\phi'(v)]$ in the Euclidean plane. Function ϕ' being bijective, this set of tangents is in one-to-one correspondence with the set of slopes $\{\phi'(v), v \in \mathbb{R}^+\} = (0, M]$. The tangent of slope $z \in (0, M]$ is defined as:

$$y = zx + \underbrace{[\phi \circ (\phi')^{-1}(z) - z(\phi')^{-1}(z)]}_{\triangleq \psi(z)}.$$

Since ϕ is strictly concave, all these tangent lines are above the graph of ϕ (Fig. 17), i.e,

$$zv + \psi(z) \geq \phi(v),\ \forall v \in \mathbb{R}^+,\ \forall z \in (0, M]\ ,$$

the inequality being strict, except for $z = \phi'(v)$. Hence:

$$\forall v \in \mathbb{R}^+,\ \left\{ \begin{array}{l} \phi(v) = \min_{z\in(0,M]} zv + \psi(z), \\ \phi'(v) = \arg\min_{z\in(0,M]} zv + \psi(z)\ , \end{array} \right.$$

which is equivalent to:

$$\forall u \in \mathbb{R},\ \left\{ \begin{array}{l} \rho(u) = \phi(u^2) = \min_{z\in(0,M]} zu^2 + \psi(z), \\ \frac{\rho'(u)}{2u} = \phi'(u^2) = \arg\min_{z\in(0,M]} zu^2 + \psi(z). \end{array} \right.$$

Let's come back to function ψ. It is continuously differentiable on $(0, M]$, with $\psi'(z) = -(\phi')^{-1}(z) < 0$. Therefore it is strictly decreasing, with $\psi(M) = \rho(0)$ (usually zero), and ψ tends to some finite or infinite limit when $z \to 0+$.

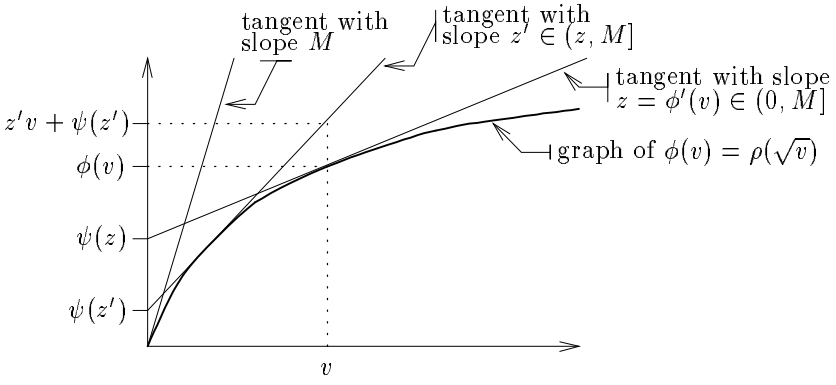

Figure 17: Graph of ϕ as the inferior envelope of its set of tangent lines (parameterized by their slope $z \in (0, M]$) – graphic illustration of $\psi(z)$ as the y-intercept of the tangent with slope z.

B Minimization of $\mathbb{H}$ with respect to β_{sr}s

Let $<s, r>$ be a pair of neighboring sites. The part of $\mathbb{H}$ which actually depends on β_{sr} is:

$$\beta_{sr}\left(M_2\|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{|\mathcal{C}_{\partial R}|}\mathbf{1}_{\partial R}(s,r) - \frac{\mu_2'}{|\mathcal{C}_R|}\mathbf{1}_R(s,r)\right) + \psi_2(M_2\beta_{sr})\ . \tag{42}$$

Therefore, if

$$M_2\|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{|\mathcal{C}_{\partial R}|}\mathbf{1}_{\partial R}(s,r) - \frac{\mu_2'}{|\mathcal{C}_R|}\mathbf{1}_R(s,r) < 0, \tag{43}$$

which means that $<s,r>\subset R$ and $\|\Delta \boldsymbol{w}_{sr}\|^2 < \frac{\mu_2'}{M_2|\mathcal{C}_R|}$, then $\mathbb{H}$ is a decreasing function of β_{sr} (since ψ_2 is decreasing) whose minimizer in $(0,1]$ is $1 = \frac{\phi_2'(0)}{M_2}$. In other cases, the minimizer zeroes the partial derivative:

$$\begin{aligned}
&\frac{\partial \mathbb{H}}{\partial \beta_{sr}}(\hat{\beta}_{sr}) = 0 \\
&\Leftrightarrow M_2\|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{|\mathcal{C}_{\partial R}|}\mathbf{1}_{\partial R}(s,r) - \frac{\mu_2'}{|\mathcal{C}_R|}\mathbf{1}_R(s,r) + M_2\psi_2'(M_2\hat{\beta}_{sr}) = 0 \qquad (44)\\
&\Leftrightarrow \hat{\beta}_{sr} = \frac{1}{M_2}\phi_2'\left[\|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{M_2|\mathcal{C}_{\partial R}|}\mathbf{1}_{\partial R}(s,r) - \frac{\mu_2'}{M_2|\mathcal{C}_R|}\mathbf{1}_R(s,r)\right] \qquad (45)
\end{aligned}$$

since $\psi_2' = -(\phi_2')^{-1}$. The optimal weight computation rule is then:

- if the site pair is outside R ($<s,r>\subset \bar{R}$):

$$\hat{\beta}_{sr} = \frac{1}{M_2}\phi_2'\left[\|\Delta \boldsymbol{w}_{sr}\|^2\right];$$

- if the site pair straddles the "border" of R ($<s,r>\in R\times\bar{R}$):

$$\hat{\beta}_{sr} = \frac{1}{M_2}\phi_2'\left[\|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{M_2|\mathcal{C}_{\partial R}|}\right];$$

- if the site pair is inside R ($<s,r>\subset R$):

$$\hat{\beta}_{sr} = \frac{1}{M_2}\phi_2'\left(\left[\|\Delta \boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{M_2|\mathcal{C}_R|}\right]^+\right) = \begin{cases} \frac{1}{M_2}\phi_2'\left[\|\Delta \boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{M_2|\mathcal{C}_R|}\right] & \text{if } \|\Delta \boldsymbol{w}_{sr}\|^2 > \frac{\mu_2'}{M_2|\mathcal{C}_R|}\,, \\ 1 \text{ otherwise.} \end{cases}$$

Also, notice that in case (43) does not hold, it comes from (44) that:

$$\begin{aligned}
&\hat{\beta}_{sr}\left(M_2\|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{|\mathcal{C}_{\partial R}|}\mathbf{1}_{\partial R}(s,r) - \frac{\mu_2'}{|\mathcal{C}_R|}\mathbf{1}_R(s,r)\right) + \psi_2(M_2\hat{\beta}_{sr}) \\
&\quad = -\hat{\beta}_{sr}M_2\psi_2'(M_2\hat{\beta}_{sr}) + \psi_2(M_2\hat{\beta}_{sr}) \\
&\quad = M_2\hat{\beta}_{sr}(\phi_2')^{-1}(M_2\hat{\beta}_{sr}) + \psi_2(M_2\hat{\beta}_{sr}) \\
&\quad = \phi_2 \circ (\phi_2')^{-1}(M_2\hat{\beta}_{sr})\,, \qquad (46)
\end{aligned}$$

since $\psi_2(z) = \phi_2 \circ (\phi_2')^{-1}(z) - z(\phi_2')^{-1}(z)$ by definition. If $\phi_2(v) = 1 - \exp\{-M_2 v\}$ (Leclerc's estimator), the above energetic contribution becomes:

$$\hat{\beta}_{sr}\left(M_2\|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{|\mathcal{C}_{\partial R}|}\mathbf{1}_{\partial R}(s,r) - \frac{\mu_2'}{|\mathcal{C}_R|}\mathbf{1}_R(s,r)\right) + \psi_2(M_2\hat{\beta}_{sr}) = 1 - \hat{\beta}_{sr}. \qquad (47)$$

C Computation and approximation of $\Delta E(R)$ for Leclerc's estimator

Suppose $\boldsymbol{w}$ is fixed. Let R be some segment, and β the optimal set of discontinuity weights associated to it according to the computation previously derived. The part of energy concerned by the segment updating process is the one which involves the segment and/or these weights:

$$\begin{aligned}
E(R) &= E_{prior}(R) \\
&\quad + \sum_{<s,r>\subset\bar{R}} M_2\hat{\beta}^0_{sr}\|\Delta\boldsymbol{w}_{sr}\|^2 + \psi_2(M_2\hat{\beta}^0_{sr}) \\
&\quad + \sum_{<s,r>\in\bar{R}\times R} K_{\partial R}\hat{\beta}^0_{sr}\left(M_2\|\Delta\boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{|\mathcal{C}_{\partial R}|}\right) + \psi_2(M_2K_{\partial R}\hat{\beta}^0_{sr}) \\
&\quad + \sum_{<s,r>\subset R:\|\Delta\boldsymbol{w}_{sr}\|^2\geq\frac{\mu_2'}{M_2|\mathcal{C}_R|}} K_R\hat{\beta}^0_{sr}\left(M_2\|\Delta\boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{|\mathcal{C}_R|}\right) + \psi_2(M_2K_R\hat{\beta}^0_{sr}) \\
&\quad + \sum_{<s,r>\subset R:\|\Delta\boldsymbol{w}_{sr}\|^2<\frac{\mu_2'}{M_2|\mathcal{C}_R|}} M_2\|\Delta\boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{|\mathcal{C}_R|} + \underbrace{\psi_2(M_2)}_{=0}.
\end{aligned}$$

Using (47) and (39-40-41), it reduces to:

$$\begin{aligned}
E(R) &= E_{prior}(R) + \sum_{<s,r>\subset\bar{R}}(1-\hat{\beta}^0_{sr}) + \sum_{<s,r>\in\bar{R}\times R}(1-K_{\partial R}\hat{\beta}^0_{sr}) \\
&\quad + \sum_{<s,r>\subset R:\|\Delta\boldsymbol{w}_{sr}\|^2\geq\frac{\mu_2'}{M_2|\mathcal{C}_R|}}(1-K_R\hat{\beta}^0_{sr}) + \sum_{<s,r>\subset R:\|\Delta\boldsymbol{w}_{sr}\|^2<\frac{\mu_2'}{M_2|\mathcal{C}_R|}} M_2\|\Delta\boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{|\mathcal{C}_R|}.
\end{aligned}$$

In the latest term of the left-hand side, if $M_2\|\Delta\boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{|\mathcal{C}_R|} << 0$, then a first order Taylor expansion yields:

$$M_2\|\Delta\boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{|\mathcal{C}_R|} \approx 1 - \exp\left\{-M_2\|\Delta\boldsymbol{w}_{sr}\|^2 + \frac{\mu_2'}{|\mathcal{C}_R|}\right\} = 1 - K_R\hat{\beta}^0_{sr}.$$

Then:

$$\begin{aligned}
E(R) &\approx E_{prior}(R) - \sum_{<s,r>\subset\bar{R}}\hat{\beta}^0_{sr} - K_{\partial R}\sum_{<s,r>\in\bar{R}\times R}\hat{\beta}^0_{sr} - K_R\sum_{<s,r>\subset R}\hat{\beta}^0_{sr} + |\mathcal{C}| \\
&= \underbrace{E_{prior}(R) + (1-K_{\partial R})\sum_{<s,r>\in\bar{R}\times R}\hat{\beta}^0_{sr} + (1-K_R)\sum_{<s,r>\subset R}\hat{\beta}^0_{sr}}_{\triangleq\tilde{E}(R)} + |\mathcal{C}| - \sum_{<s,r>}\hat{\beta}^0_{sr}.
\end{aligned}$$

The velocity field $\boldsymbol{w}$ being fixed (and therefore such is $\hat{\beta}^0$), the energy comparison of two segments R and R' is then approximately driven by:

$$E(R') - E(R) \approx \tilde{E}(R') - \tilde{E}(R),$$

which is very easy to compute. Especially, if R and R' have the same shape, the only terms changing from $\tilde{E}(R)$ to $\tilde{E}(R')$ are $\sum_{<s,r>\subset R}\hat{\beta}^0_{sr}$ and $\sum_{<s,r>\in\bar{R}\times R}\hat{\beta}^0_{sr}$.

References

[1] G. AUBERT, M. BARLAUD, L. BLANC-FÉRAUD, and P. CHARBONNIER. Deterministic edge-preserving regularization in computed imaging. Technical Report 94-01, Université de Nice, Laboratoire I3S, January 1994.

[2] J. BARRON, D. FLEET, and S. BEAUCHEMIN. Performance of optical flow techniques. *Int. J. Computer Vision*, 12(1):43–77, 1994.

[3] B. BASCLE and R. DERICHE. Region tracking through image sequences. Technical report, INRIA, December 1994.

[4] B. BASCLE and R. DERICHE. Region tracking through image sequences. In *Proc. Int. Conf. Computer Vision*, pages 302–307, Cambridge, USA, June 1995.

[5] M. BLACK. *Robust incremental optical flow.* PhD thesis, Yale University, 1992.

[6] M. BLACK. Recursive non-linear estimation of discontinuous flow fields. In *Proc. Europ. Conf. Computer Vision*, pages 138–145, Stockholm, Sweden, 1994.

[7] M. BLACK and P. ANANDAN. Robust incremental optical flow. In *Proc. Conf. Comp. Vision Pattern Rec.*, 1992.

[8] A. BLAKE, R. CURWEN, and A. ZISSERMAN. A framework for spatiotemporal control in the tracking of visual contours. *Int. J. Computer Vision*, 11(2):127–145, 1993.

[9] A. BLAKE and A. ZISSERMAN. *Visual reconstruction.* The MIT Press, Cambridge, USA, 1987.

[10] L. BLANC-FÉRAUD, M. BARLAUD, and T. GAIDON. Motion estimation involving discontinuities in a multiresolution scheme. *Optical engineering*, 32(7):1475–1482, July 1993.

[11] P. BOUTHEMY and E. FRANÇOIS. Motion segmentation and qualitative dynamic scene analysis from an image sequence. *Int. J. Computer Vision*, 10(2):157–182, 1993.

[12] R. CHELLAPPA and A.K. JAIN, editors. *Markov random fields: theory and applications.* Academic Press, Boston, 1993.

[13] I. COHEN. Nonlinear variational method for optical flow computation. In *Proc. Scand. Conf. Image Analysis*, pages 1:523–530, Tromso, Norway, May 1993.

[14] L. COHEN and I. COHEN. Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Trans. Pattern Anal. Machine Intell.*, 15(11):1131–1147, 1993.

[15] R. DERICHE, P. KORNPROBST, and G. AUBERT. Optical flow estimation while preserving its discontinuities: a variational approach. In *Proc. Asian Conf. Computer Vision*, volume 1, pages 290–295, Singapore, December 1995.

[16] W. ENKELMANN. Investigation of multigrid algorithms for the estimation of optical flow fields in image sequences. *Comp. Vision Graph. and Image Proces.*, 43:150–177, 1988.

[17] K. FUJIMURA, N. YOKOYA, and K. YAMAMOTO. Motion tracking of deformable objects by active contour models using multiscale dynamic programming. *J. Visual Comm. Image Representation*, 4(4):382–391, 1993.

[18] D. GEMAN and G. REYNOLDS. Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Machine Intell.*, 14(3):367–383, 1992.

[19] S. GEMAN and D. GEMAN. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, 6(6):721–741, 1984.

[20] S. GEMAN and D. E. McCLURE. Statistical methods for tomographic image reconstruction. In *Bull. ISI, Proc. 46th Session Int. Statistical Institute*, volume 52, 1987.

[21] U. GRENANDER. *General Pattern theory*. Oxford Univ. Press, 1993.

[22] U. GRENANDER and K. MANBECK. A stochastic shape and color model for defect detection in potatoes. *Applied Statistics*, 2(2):131–151, 1993.

[23] U. GRENANDER and M.I. MILLER. Representations of knowledge in complex systems. *J. Royal Statis. Soc. B*, 56(3), 1994.

[24] F. HEITZ and P. BOUTHEMY. Multimodal estimation of discontinuous optical flow using Markov random fields. *IEEE Trans. Pattern Anal. Machine Intell.*, 15(12):1217–1232, 1993.

[25] F. HEITZ, P. PÉREZ, and P. BOUTHEMY. Multiscale minimization of global energy functions in some visual recovery problems. *CVGIP : Image Understanding*, 59(1):125–134, 1994.

[26] P. HOLLAND and R. WELSCH. Robust regression using iteratively reweighted least-squares. *Commun. Statis.-Theor. Meth.*, A6(9):813–827, 1977.

[27] B. HORN and B. SCHUNCK. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[28] P. HUBER. *Robust Statistics*. John Wiley & Sons, 1981.

[29] M. KASS, A. WITKIN, and D. TERZOPOULOS. Snakes : active contour models. *Int. J. Computer Vision*, 1(4):321–331, 1988.

[30] C. KERVRANN and F. HEITZ. A hierarchical statistical framework for the segmentation of deformable objects in image sequences. In *Proc. Comp. Vision Pat. Recog.*, pages 724–728, Seattle, USA, June 1994.

[31] Y. LECLERC. Constructing simple stable descriptions for image partitioning. *Int. J. Computer Vision*, 3:73–102, 1989.

[32] F. LEYMARIE and M.D. LEVINE. Tracking deformable objects in the plane using an active contour model. *IEEE Trans. Pattern Anal. Machine Intell.*, 15(6):617–634, 1993.

[33] S. Z. LI. On discontinuity-adaptative smoothness priors in computer vision. *IEEE Trans. Pattern Anal. Machine Intell.*, 17(6):576–586, 1995.

[34] E. MÉMIN, F. HEITZ, and F. CHAROT. Efficient parallel non-linear multigrid relaxation algorithms for low-level vision applications. *J. Parallel and Distributed Computing*, 29(1):96–103, 1995.

[35] F. MEYER and P. BOUTHEMY. Region-based tracking using affine motion models in long image sequences. *CVGIP: Image Understanding*, 60(2), 1994.

[36] D. MUMFORD and J. SHAH. Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. Pure and Appl. Math.*, 42:577–685, 1989.

[37] P. NESI. Variational approach to optical flow estimation managing discontinuities. *Image and Vision Computing*, 11(7):419–439, 1993.

[38] P. PÉREZ and B. GIDAS. Motion segmentation and tracking using deformable templates. In *Proc. Int. Conf. Image Processing*, Austin, November 1994.

[39] T. POGGIO, V. TORRE, and C. KOCH. Computational vision and regularization theory. In M.A. Fischler and O. Firschein, editors, *Readings in Computer Vision : Issues, Problems, Principles and Paradigms*, pages 638–643. Morgan Kaufmann, 1987.

[40] P. PROESMANS, E. PAUWELS, and L. van GOOL. Coupled geometry-driven diffusion equations for low-level vision. In B. ter HAAR ROMENY, editor, *Geometry-driven diffusion in computer vision*. Kluwer Academic Publishers, Dordrecht, 1995.

[41] P. PROESMANS, L. van GOOL, E. PAUWELS, and A. OOSTERLINCK. Determination of optical flow and its discontinuities using non-linear diffusion. In *Proc. Europ. Conf. Computer Vision*, volume 2, pages 295–304, Stokholm, Sweden, 1994.

[42] C. SCHNORR and W. PECKARD. Motion-based identification of deformable templates. In V. Hlavac and R. Sara, editors, *Proc. Int. Conf. Computer Analysis of Images and Patterns*, number 970 in Lecture notes in computer science, Prague, September 1995.

[43] C. STILLER and B. HURTGEN. Combined displacement estimation and segmentation in image sequences. In *SPIE/EUROPTO Video Comm. and PACS for Medical Appl.*, pages 276–287, Berlin, Germany, April 1993.

[44] R. SZELISKI and J. COUGHLAN. Spline-based image registration. Technical Report CRL 94/1, Digital Equipment Corporation, April 1994.

[45] T. VIÉVILLE and O. FAUGERAS. Robust and fast computation of unbiased intensity derivatives in images. In *Proc. Europ. Conf. Computer Vision*, pages 203–212, S. Margherita Ligure, May 1992.

Unit´e de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit´e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit´e de recherche INRIA Rhône-Alpes, 46 avenue F´elix Viallet, 38031 GRENOBLE Cedex 1
Unit´e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit´e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399